



PCT/AU98/00743

REC'D 29 SEP 1998

WIPO PCT

Patent Office  
Canberra

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PO 9108 for a patent by COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION and THE AUSTRALIAN NATIONAL UNIVERSITY filed on 12 September 1997.

I further certify that the annexed specification is not, as yet, open to public inspection.

**PRIORITY DOCUMENT**

WITNESS my hand this Seventeenth  
day of September 1998

KIM MARSHALL  
MANAGER EXAMINATION SUPPORT AND  
SALES



AUSTRALIA  
Patents Act 1990

**PROVISIONAL SPECIFICATION**

**Applicant(s):** COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH  
ORGANISATION  
and  
THE AUSTRALIAN NATIONAL UNIVERSITY

**Invention Title:** REGULATION OF GENE EXPRESSION IN PLANTS

The invention is described in the following statement:

## REGULATION OF GENE EXPRESSION IN PLANTS

This invention relates to methods of modulating the expression of desired genes in plants, and to DNA sequences and genetic constructs for use in these methods. In one preferred embodiment, the invention relates to methods and constructs for targeting of expression specifically to the endosperm of the seeds of cereal plants such as wheat, and for modulating the time of expression in the target tissue. This is achieved by the use of promoter sequences from enzymes of the starch biosynthetic pathway. In preferred embodiments of the invention, the sequences and/or promoters are those of starch branching enzyme I, starch branching enzyme II, soluble starch synthase, and starch debranching enzyme, all derived from *Triticum tauschii*, the D genome donor of hexaploid bread wheat.

### BACKGROUND OF THE INVENTION

Starch is an important constituent of the cereal grains and of flours, accounting for about 65-67% of the weight of the grain at maturity. It is produced in the amyloplast of the grain endosperm by the concerted action of a number of enzymes, including ADP-Glucose pyrophosphorylase (EC 2.7.7.27), starch synthases (EC 2.4.1.21), branching enzymes (EC 2.4.1.18) and debranching enzymes (EC 3.2.1.41 and EC 3.2.1.68) (Ball et al, 1996; Martin and Smith, 1995; Morell et al, 1995). Some of the proteins involved in the synthesis of starch can be recovered from the starch granule (Denyer et al, 1995; Rahman et al, 1995).

Most wheat cultivars normally produce starch containing 25% amylose and 75% amylopectin. Amylose is composed of large linear chains of  $\alpha$  (1-4) linked  $\alpha$ -D-glucopyranosyl residues, whereas amylopectin is a branching form of  $\alpha$ -glycan linked by  $\alpha$  (1-6) linkages. The ratio of amylose and amylopectin, the branch chain length and the

number of branch chains of amylopectin are the major factors which determine the properties of wheat starch.

Wheat contains three sets of chromosomes (A, B and D) in its very large genome of about  $10^{10}$  bp. The  
5 donor of the D genome to wheat is *Triticum tauschii*, and by using a suitable accession of this species the genes from the D genome can be studied separately (Lagudah et al, 1991).

Starch with various properties has been widely  
10 used in industry, food science and medical science. High amylose wheat can be used for the plastic substitutes and in paper manufacture to protect the environment; in health foods to reduce bowel cancer and heart disease; and in  
15 sport food to improve the athletes' performance. High amylopectin wheat may be suitable for Japanese noodles and is used as a thickener in the food industry.

There is comparatively little variation in starch structure found in wheat varieties, because the hexaploid nature of wheat prevents mutations from being readily  
20 identified. Dramatic alterations in starch structure are expected to require the combination of homozygous recessive alleles from each of the 3 wheat genomes, A, B and D. This requirement renders the probability of finding such mutants in natural or mutagenised populations very low. Variation  
25 in wheat starch is desirable in order to enable better tailoring of wheat starches for processing and end-user requirements.

Key commercial targets for the manipulation of starch biosynthesis are:

30 1. "Waxy" wheats in which amylose content is decreased to insignificant levels. This outcome is expected to be obtained by eliminating granule-bound starch synthase activity.

2. High amylose wheats, expected to be  
35 obtained by suppressing starch branching enzyme-II activity.



3. Wheats which continue to synthesise starch at elevated temperatures, expected to be obtained by identifying or introducing a gene encoding a heat stable soluble starch synthase.

- 5 4. "Sugary types" of wheat which contain increased amylose content and free sugars, expected to be obtained by manipulating an isoamylase-type debranching enzyme.

There are two general strategies for obtaining wheats with altered starch structure:

- (a) using genetic engineering strategies to suppress the activity of a specific gene, or to introduce a novel gene into a wheat line; and
- (b) selecting among existing variation in wheat for missing (or "null") or altered alleles of a gene in each of the genomes of wheat, and combining these by plant breeding.

Branching enzymes are involved in the production of glucose  $\alpha$  (1,6) branches. Of the two main constituents of starch, amylose is essentially linear, but amylopectin is highly branched; thus branching enzymes are thought to be directly involved in the synthesis of amylopectin but not amylose. There are two types of branching enzymes in plants, starch branching enzyme I (SBE I) and starch branching enzyme II (SBE II), and both are about 85 kDa in size. At the nucleic acid level there is about 65% sequence identity between types I and II in the central portion of the molecules; the sequence identity between SBE I from different cereals is about 85% overall (Burton et al, 1995; Morell et al, 1995).

In cereals, SBE I genes have so far been reported only for rice (Kawasaki et al, 1991) and wheat (Rahman et al, 1997). A cDNA sequence for wheat SBE I is available on the Genbank database. As far as we are aware, no promoter sequence for wheat SBE I has been reported. We have characterised an SBE I gene (called *wSBE I-D2*) from *Triticum tauschii* (the donor of the D genome to wheat),

that encoded a novel protein sequence in that it was missing approximately 65 amino acids at the C-terminal end, and appeared not to contain some of the conserved amino acid motifs characteristic of this class of enzyme

5 (Svensson, 1994). Although *wSBE I-D2* was expressed as mRNA, no corresponding protein has yet been found in our analysis of SBE I isoforms from the endosperm, and thus it is possible that this gene is a transcribed pseudogene.

10 Genes for SBE II are less well characterised; the only one for cereals which has been fully characterised is that of rice (Kawasaki *et al*, 1993). A cDNA sequence for SBE II from wheat is available on the Genbank database; although the sequences are very similar to those reported herein there are differences near the N-terminal of the  
15 protein, which specifies its intracellular location. No promoter sequences have been reported, as far as we are aware.

Wheat granule-bound starch synthase (GBSS) is responsible for amylose synthesis, while wheat branching  
20 enzymes together with soluble starch synthases are considered to be directly involved in amylopectin biosynthesis. A number of isoforms of soluble and granule-bound starch synthases have been identified in developing wheat endosperm (Denyer *et al*, 1995). There are three  
25 distinct isoforms of starch synthases, 60 kDa, 75 or 77 kDa and 100-105 kDa, which exist in the starch granules (Denyer *et al*, 1995; Rahman *et al*, 1995). The 60 kDa GBSS is the product of the *wx* gene. The 75 or 77 kDa protein is a wheat soluble starch synthase (SSS) which is present in  
30 both the soluble fraction and the starch granule-bound fraction of the endosperm. However, the 100-105 kDa proteins, which are another type of soluble starch synthase, are located only in starch granules (Denyer *et al*, 1995; Rahman *et al*, 1995). To our knowledge there has  
35 been no report of any complete wheat SSS sequence, either at the protein or the nucleotide level.

Both cDNA and genomic DNA encoding a soluble starch synthase I of rice have been cloned and analysed (Baba et al, 1993; Tanaka et al, 1995). The cDNAs encoding potato soluble starch synthase SSII and SSIII and pea soluble starch synthase SSII have also been reported (Edwards et al, 1995; Marshall et al, 1996; Gernot et al, 1996; Dry et al, 1992). However, corresponding cDNA sequences for wheat have hitherto not been available, and the genes for other enzymes involved in starch biosynthesis have not previously been isolated.

Approach (b) referred to above has been demonstrated for the gene granule bound starch synthase. Null alleles on chromosomes 7A, 7D and 4A were identified by the analysis of GBSS protein bands by electrophoresis, and combined by plant breeding to produce a wheat line containing no GBSS, and no amylose. Subsequently, PCR-based DNA markers have been identified, which also identify null alleles for the GBSS loci on each of the three wheat genomes.

#### SUMMARY OF THE INVENTION

In this application we report the isolation and identification of novel genes from *T. tauschii*, the D-genome donor of wheat, that encode SBE I, SBE II, a 75 kDa SSS, and an isoamylase-type debranching enzyme (DBE). Because of the very close relationship between *T. tauschii* and wheat, as discussed above, results obtained with *T. tauschii* can be directly applied to wheat with little if any modification. Such modification as may be required represents routine trial and error experimentation. Sequences from these genes can be used as probes to identify null or altered alleles in wheat, which can then be used in plant breeding programmes to provide modifications of starch characteristics. The novel sequences of the invention can be used in genetic engineering strategies or to introduce a desired gene into a host plant, to provide antisense sequences for

suppression of one or more specific genes in a host plant, in order to modify the characteristics of starch produced by the plant.

In its most general aspect, the invention  
5 provides a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, said enzyme being selected from the group consisting of starch  
branching enzyme I, starch branching enzyme II, starch  
soluble synthase, and debranching enzyme, with the proviso  
10 that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.

Preferably the nucleic acid sequence is a DNA sequence, and may be genomic DNA or cDNA. Preferably the sequence is one which is functional in wheat. More  
15 preferably the sequence is derived from *Triticum* species, most preferably *Triticum tauschii*.

Where the sequence encodes soluble starch synthase, preferably the sequence encodes the 75 kD soluble starch synthase of wheat.

20 Biologically-active untranslated control sequences of genomic DNA are also within the scope of the invention. Thus the invention also provides the promoter of an enzyme as defined above.

In a preferred embodiment of this aspect of the  
25 invention, there is provided a genetic construct comprising a nucleic acid sequence of the invention, a biologically-active fragment thereof, or a fragment thereof encoding a biologically-active fragment of an enzyme as defined above, operably linked to one or more nucleic acid sequences  
30 facilitating expression of said enzyme in a plant, preferably a cereal plant. The construct may be a plasmid or a vector, preferably one suitable for use in transformation of plant. Such a suitable vector is a bacterium of the genus *Agrobacterium*, preferably  
35 *Agrobacterium tumefaciens*. Methods of transforming cereal plants using *Agrobacterium tumefaciens* are known; see for

example Australian Patent No. 667939 by Japan Tobacco Inc., and Tingay et al (1997).

In a second aspect, the invention provides a genetic construct for targeting of a desired gene to  
5 endosperm of a cereal plant, and/or for modulating the time of expression of a desired gene in endosperm of a cereal plant, comprising one or more promoter sequences selected from SBE I promoter, SBE II promoter, SSS promoter, and DBE promoter, operatively linked to a nucleic acid sequence  
10 encoding a desired protein, and optionally also operatively linked to one or more additional targeting sequences and/or one or more 3' untranslated sequences.

The nucleic acid encoding the desired protein may be in either the sense orientation or in the antisense  
15 orientation. Preferably the desired protein is an enzyme of the starch biosynthetic pathway. For example, the antisense sequences of GBSS, starch debranching enzyme, SBE II, low molecular weight glutenin, or grain softness protein I, may be used. Preferred sequences for use in  
20 sense orientation include those of bacterial isoamylase, bacterial glycogen synthase, or wheat high molecular weight glutenin Bx17. It is contemplated that any desired protein which is encoded by a gene which is capable of being expressed in the endosperm of a cereal plant is suitable  
25 for use in the invention.

In a third aspect, the invention provides a method of modifying the characteristics of starch produced by a plant, comprising the step of:

(a) introducing a gene encoding a desired  
30 enzyme of the starch biosynthetic pathway into a host plant, and/or

(b) introducing an anti-sense nucleic acid sequence directed to a gene encoding an enzyme of the starch biosynthetic pathway into a host plant,  
35 wherein said enzymes are as defined above.

Where both steps (a) and (b) are used, the enzymes in the two steps are different.

Preferably the plant is a cereal plant, more preferably wheat or barley.

As is well known in the art, anti-sense sequences can be used to suppress expression of the protein to which the anti-sense sequence is complementary. It would be evident to the person skilled in the art that different combinations of sense and anti-sense sequences may be chosen so as to effect a variety of different modifications of the characteristics of the starch produced by the plant.

In a fourth aspect, the invention provides a method of targeting expression of a desired gene to the endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to the invention.

According to a fifth aspect, the invention provides a method of modulating the time of expression of a desired gene in endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to the second aspect of the invention.

Where expression at an early stage following anthesis is desired, the construct preferably comprises the SBE II promoter. Where expression at a later stage following anthesis is desired, the construct preferably comprises the SBE I promoter.

In a sixth aspect, the invention provides a method of identifying a null or altered allele encoding an enzyme of the starch biosynthetic pathway, comprising the steps of subjecting DNA from a plant suspected to possess such an allele to a DNA fingerprinting assay, wherein DNA probes used in the assay comprise one or more of the nucleic acid sequences of the invention. The nucleic acid sequence may be a genomic DNA or a cDNA, and may comprise the full-length coding sequence or a fragment thereof.

DNA fingerprinting methods are well known in the art, and any suitable technique may be used.

While the invention is described in detail in relation to wheat, it will be clearly understood that it is

also applicable to other cereal plants of the family Gramineae, such as maize, barley and rice.

Methods for transformation of monocotyledonous plants such as wheat, maize, barley and rice and for  
5 regeneration of plants from protoplasts or immature plant embryos are well known in the art. See for example Lazzeri et al, 1991; Jahne et al, 1991 and Wan and Lenaux, 1994 for barley; Wirtzens et al, 1997; Tingay et al, 1997; Canadian  
10 Patent Application No. 2092588 by Nehra; Australian Patent Application No. 61781/94 by National Research Council of Canada, and Australian Patent No. 667939 by Japan Tobacco Co.

The sequences of ADP glucose pyrophosphorylase from barley (Australian Patent Application No. 65392/94),  
15 starch debranching enzyme and its promoter from rice (Japanese Patent Publication No. Kokai 6261787 and Japanese Patent Publication No. Kokai 5317057), and starch debranching enzyme from spinach and potato (Australian Patent Application No. 44333/96) are all known.

20

#### Detailed Description of the Drawings

The invention will be described in detail by reference only to the following non-limiting examples and to the figures.

25

Figure 1 shows the hybridisation of genomic clones isolated from *T. tauschii*.

DNA was extracted from the different clones, digested with *Bam*HI and hybridised with the 5' end of the maize SBE I cDNA. Lanes 1, 2, 3 and 4 correspond to DNA  
30 from clones  $\lambda$ E1,  $\lambda$ E2,  $\lambda$ E6 and  $\lambda$ E7 respectively. Note that clones  $\lambda$ E1 and  $\lambda$ E2 give identical patterns, the SBE I gene in  $\lambda$ E6 is a truncated form of that in  $\lambda$ E1, and  $\lambda$ E7 gives a clearly different pattern.

Figure 2 shows the hybridisation of DNA from  
35 *T. tauschii*.

DNA from *T. tauschii* was digested with *Bam*HI and the hybridisation pattern compared with DNA from  $\lambda$ E1 and

$\lambda$ E7 digested with the same enzyme. Fragment E1.1 (see Figure 3) from  $\lambda$ E1 was used as the probe; it contains some sequences that are over 80% identical to sequences in E7.8. Approximately 25  $\mu$ g of *T. tauschii* DNA was electrophoresed in lane 1, and 200 pg each of  $\lambda$ E1 and  $\lambda$ E7 in lanes 2 and 3, respectively.

Figure 3 shows the restriction maps of clone  $\lambda$ E1 and  $\lambda$ E7. The fragments obtained with EcoRI and BamHI are indicated. The fragments sequenced from  $\lambda$ E1 are E1.1, E1.2, a part of E1.7 and a part of E1.5.

Figure 4 shows the nucleotide sequence of part of wSBE I-D4, and the deduced amino acid sequence and N-terminal sequence of SBE I (Morell *et al*, 1997).

Figure 5 shows the hybridisation of SBE I genomic clones with the following probes,

A. wSBE I-D45 (derived from the 5' end of the gene and including sequence from fragments E1.1 and E1.7), and

B. wSBE I-D43, (derived from the 3' end of the gene and containing sequences from fragment E1.5).

Clones  $\lambda$ E7 and  $\lambda$ E22 do not hybridise to either of the probes, but do hybridise with a probe from E1.1. Clone  $\lambda$ E30 contains a sequence unrelated to SBE I.

Figure 6 shows the intron-exon structure of wSBE I-D4 compared to the corresponding structures of rice SBE I (Kawasaki *et al*, 1993) and wSBE I-D2 (Rahman *et al*, 1997). The intron-exon structure of wSBE I-D4 is deduced by comparison with the major species of wSBE I-D4 cDNA present in the endosperm.

The dark rectangles correspond to exons and the light rectangles correspond to introns. The bars above the structures indicate the percentage identity in sequence between the indicated exons and introns of the relevant genes. Note that intron 2 shares no significant sequence identity and is not indicated.

Figure 7 shows the alignment of cDNA clones to obtain the sequence represented by wSBE I-D4 cDNA. BED4



and BED5 were obtained from screening the cDNA library with maize BEI (Baba *et al*, 1991). BED1, 2 and 3 were obtained by RT-PCR using defined primers.

Figure 8 shows the amino acid sequence of SBE I as deduced from the sequence of wSBE I-D4 cDNA. The N-terminal sequence of SBE I (Morell *et al*, 1997) is in bold, and residues considered by Svensson (1994) to be invariant in the  $\alpha$ -amylase family are underlined.

Figure 9 shows the comparison of deduced amino acid sequence of wSBE I-D4 cDNA with the deduced amino acid sequence of rice SBE I (Nakamura *et al*, 1992), maize SBE I (Baba *et al*, 1991), wSBE I-D2 type cDNA (Rahman *et al*, 1997), pea SBE II (equivalent to maize SBE I, Burton *et al*, 1995), and potato SBE I (Cangiano *et al*, 1993). Residues present in at least three of the sequences are identified in the consensus sequence in capitals.

Figure 10 shows the expression of SBE I type sequences during endosperm development. The probe used was wSBE I-D43C, corresponding to the untranslated 3' end of wSBE I-D4 cDNA. There is no hybridisation to RNA extracted from leaves or florets prior to anthesis.

Figure 11 shows the comparison of wSBE I-D4 and rice SBE I genomic sequence (Kawasaki *et al*, 1993; Accession Number D10838) using the programs Compares and DotPlot (Devereaux *et al*, 1984). The programs used a window of 21 bases with a stringency of 14 to register a dot.

Figure 12 shows the hybridisation of wheat DNA from chromosome-engineered lines using the following probes:

- A. wSBE I-D45 (from the 5' end of the gene),
- B. wSBE I-D43 (from the 3' end of the gene),

and

- C. wSBE I-D4R (repetitive sequence approximately 600 bp 3' to the end of wSBE I-D4 sequence.
- N7AT7B, no 7A chromosome, four copies of 7B chromosome; N7BT7D, no 7B chromosome, four copies of 7D

chromosome; NTDT7A, no 7D chromosome, four copies of 7A chromosome. The chromosomal origin of hybridising bands is indicated.

Figure 13 shows the entire sequence of the  
5 wSBE I-D4 gene. The promoter sequences is given in (a) up to the first translated amino acid. The coding sequence of the gene is given in (b), with about 50 bases of the promoter sequence.

Figure 14 shows the sequence of wSBE I-D43C,  
10 representing the 3' untranslated region of wSBE I-D4cDNA.

Figure 15 shows the hybridisation of genomic clones F1, F2, F5 and F6 with the entire SBE-9 sequence. The DNA from the clones was purified and digested with either BamHI or EcoRI, separated on agarose, blotted onto  
15 nitrocellulose and hybridised with labelled SBE-9 (a SBE II type cDNA). The pattern of hybridising bands is different in the four isolates.

Figure 16a shows the N-terminal sequence of SBE II from wheat as in Morell *et al*, (1997).

20 Figure 16b shows the deduced amino acid sequence from part of wSBE II-D1 that encodes the N-terminal sequence as described in Morell *et al*, (1997).

Figure 16c shows the deduced amino acid sequence from SBE-9 (a SBE II type cDNA).

25 Figure 17 shows the deduced exon-intron structure for a part of wSBE II-D1. The scale is marked in bases. The dark rectangles are exons.

Figure 18 shows the hybridisation of DNA from chromosome engineered lines of wheat (cultivar Chinese  
30 Spring) with a probe from nucleotides 550-850 from SBE-9. The band of approximately 2.2 kb are missing in the line in which missing chromosome 2D is absent.

T2BN2A: four copies of chromosome 2B, no copies of chromosome 2A;

35 T2AN2B: four copies of chromosome 2A, no copies of chromosome 2B;

T2AN2D: four copies of chromosome 2A, no copies of chromosome 2D.

Figure 19 shows the entire sequence of the wSBE II-D1 gene. The promoter sequence is given in (a) up to the first translated amino acid, and the coding sequence of the gene is given in (b).

Figure 20a shows the N-terminal sequence of SSS protein isolated from starch granules (Rahman et al, 1995) and deduced amino acid sequence of part of Sm2.

Figure 20b shows the nucleotide sequence of cDNA clone (sm2) for wheat soluble starch synthase.

Figure 20c shows the nucleotide sequence of genomic clone for SSS.

Figure 21 shows the deduced amino acid sequence of cDNA clone (sm2) for SSS.

Figure 22 shows the hybridisation of genomic clones sgl, 3, 4, 6 and 11 with the cDNA clone (sm2) for SSS. DNA was purified from indicated genomic clones, digested with BamHI or SacI and hybridised to sm2. Note that the hybridisation patterns for sgl, 3 and 4 are clearly different from each other.

Figure 23 shows the hybridisation of RNA from wheat tissues with the cDNA clone for SSS (sm2) as well as indicated regions of SBE II and SBE I. The tissues indicated are leaves, florets and endosperm 5-8, 10-15 and 18-22 days after anthesis (glasshouse grown material).

Figure 24 shows the hybridisation of DNA from chromosome engineered lines of wheat (cultivar Chinese Spring) digested with PvuII, with the sm2 probe.

N7AT7B: no 7A chromosome, four copies of 7B chromosome;

N7BT7D: no 7B chromosome, four copies of 7D chromosome;

N7DT7A: no 7D chromosome, four copies of 7A chromosome.

A band is missing in the N7BT7A line.

Figure 25 shows the promoter sequence of soluble

starch synthase from wheat endosperm. The sequence up to the first encoded methionine (codon ATG) is included.

Figure 26 shows the comparison of wheat and rice soluble starch synthase genomic sequences. The dark  
5 rectangles indicate exons and the light rectangles represent introns. The break indicates the area where sequencing needs to be completed.

Figure 27a shows the DNA sequence of a portion of the wheat debranching enzyme (WDBE-I) PCR product. The  
10 PCR product was generated from wheat genomic DNA (cultivar Rosella) using primers based on sequences conserved in debranching enzymes from maize and rice.

Figure 27b shows a comparison of the nucleotide sequence of wheat debranching enzyme I (WDBE-I) PCR  
15 fragment with the maize *Sugary-1* sequence.

Figure 28 shows the Southern blotting of *T. tauschii* DNA with DBE PCR product. DNA from *T. tauschii* was digested with BamHI, electrophoresed, blotted and hybridised to DBE PCR product. A band of approximately  
20 2 kb hybridised.

Figure 29 illustrates the design of 9 intron spanning BE II primer sets. Primers were based on wSBE II-D1 sequence (Figure 19) and were designed such that intron sequences in the wSBE II-D1 sequence (deduced from  
25 Figure 17) were amplified by PCR.

Figure 30 shows the results of SBE II-Intron 6 primer set on chromosome 2 nullisomic :tetrasomic lines of the wheat cultivar Chinese Spring.

30 BBD: tetra 2B nulli 2A;  
AAD: tetra 2A nulli 2B;  
AAB: tetra 2A nulli 2D;  
CS Chinese Spring normal;  
ADD: tetra 2D nulli 2B;  
ABB: tetra 2B nulli 2D;  
35 AABB: tetraploid wheat having only the A and B genomes.

The horizontal axis indicates the size of the product.

Figure 31 shows the results obtained using the SBE II-Intron 6 primer set on the wheat varieties (a) Chinese Spring and (b) Rosella.

**Example 1      Identification of Gene Encoding SBE I**  
**Construction of Genomic Library and Isolation of Clones**

The genomic library used in this study was constructed from *Triticum tauschii*, var *strangulata*, accession number CPI 100799. Of all the accessions of *T. tauschii* surveyed, the genome of CPI 100799 is the most closely related to the D genome of hexaploid wheat (Dr E. Lagudah, CSIRO Plant Industry, personal communication).

*Triticum tauschii*, var *strangulata* (CPI accession number 110799) was kindly provided by Dr E Lagudah. Leaves were isolated from plants grown in the glasshouse.

DNA was extracted from leaves of *Triticum tauschii* using published methods (Lagudah et al, 1991), partially digested with *Sau3A*, size fractionated and ligated to the arms of lambda GEM 12 (Promega). The ligated products were used to transfect the methylation tolerant strain PMC 103 (Doherty et al. 1993). A total of  $2 \times 10^6$  primary plaques were obtained with an average insert size of about 15 kb. Thus the library contains approximately 6 genomes worth of *T. tauschii* DNA. The library was amplified and stored at 4°C until required.

Positive plaques in the genomic library were selected as those hybridising with the 5' end of a maize starch branching enzyme I cDNA (Baba et al, 1991) using moderately stringent conditions as described in Rahman et al, (1997).

**Preparation of Total RNA from Wheat**

Total RNA was isolated from leaves, pre-anthesis pericarp and different developmental stages of wheat

endosperm of the cultivar, Hartog and Rosella. This material was collected from both the glasshouse and the field. The method used for RNA isolation was essentially the same as that described by Higgins *et al* (1976). RNA was then quantified by UV absorption and by separation in 1.4% agarose-formaldehyde gels which were then visualized under UV light after staining with ethidium bromide (Sambrook *et al*, 1989).

#### 10 DNA and RNA analysis

DNA was isolated and analysed using established protocols (Sambrook *et al*, 1989). DNA was extracted from wheat (cv. Chinese Spring) using published methods (Lagudah *et al*, 1991). Southern analysis was performed essentially as described by Jolly *et al* (1996). Briefly, 20 µg wheat DNA was digested, electrophoresed and transferred to a nylon membrane. Hybridisation was conducted at 42°C in 25% or 50% formamide, 2 x SSC, 6% Dextran Sulphate for 16h and the membrane was washed at 60°C in 2 x SSC for 3 x 1h unless otherwise indicated. Hybridisation was detected by autoradiography using Fuji X-Omat film.

RNA analysis was performed as follows. 10 µg of total RNA was separated in a 1.4% agarose-formaldehyde gel and transferred to a nylon Hybond N<sup>+</sup> membrane (Sambrook *et al*, 1989), and hybridized with cDNA probe at 42°C in Khandjian hybridizing buffer (Khandjian, 1989). The 3' part of wheat SBE I cDNA (designated WSBE I-D43, see Table 1) was labelled with the Rapid Multiprime DNA Probe Labelling Kit (Amersham) and used as probe. After washing at 60°C with 2 x SSC, 0.1% SDS three times, each time for about 1 to 2 hours, the membrane was visualized by overnight exposure at -80°C with X-ray film, Kodak MR.

#### Example 2

#### Frequency of Recovery of SBE I Type Clones from the Genomic Library

An estimated  $2 \times 10^6$  plaques from the amplified library were screened using an *EcoRI* fragment that

contained 1200 bp at the 5' end of maize SBE I (Baba et al, 1991) and twelve independent isolates were recovered and purified. This corresponds to the screening of somewhat fewer than the  $2 \times 10^6$  primary plaques that exist in the original library (each of which has an average insert size of 15 kb) (Maniatis et al, 1982), because the amplification may lead to the representation of some sequences more than others. Assuming that the amplified library contains approximately three genomes of *T. tauschii*, the frequency with which SBE I-positive clones were recovered suggests the existence of about 5 copies of SBE I type genes within the *T. tauschii* genome.

Digestion of DNA from the twelve independent isolates by the restriction endonuclease *Bam*HI followed by hybridisation with a maize SBE I clone, suggested that the genomic clones could be separated into two broad classes (Figure 1). One class had 10 members and a representative from this class is the clone  $\lambda$ E1 (Figure 1, lane 1);  $\lambda$ E6 (Figure 1, lane 3) is a member of this class, but is missing the 5' end of the E1-SBE I gene because the SBE I gene is at the extremity of the cloned DNA. Further hybridisation studies at high stringency with the extreme 5' and 3' regions of the SBE I gene contained in  $\lambda$ E1 suggested that the other clones contained either identical or very closely related genes.

The second family had two members, and of these clone  $\lambda$ E7 (Figure 1, lane 4) was arbitrarily selected for further study. These two members did not hybridise to probes from the extreme 5' and 3' regions of the SBE I gene that were contained in  $\lambda$ E1, indicating that they were a distinct sub-class.

The DNA from *T. tauschii* and the lambda clones  $\lambda$ E1 and  $\lambda$ E7 was digested with *Bam*HI and hybridised with fragment E1.1, as shown in Figure 2. This fragment contains sequences that are highly conserved (85% sequence identity over 0.3 kb between  $\lambda$ E1 and  $\lambda$ E7), corresponding to exons 3, 4 and 5 of the rice gene. The bands in the

genomic DNA at 0.8 kb and 1.0 kb correspond to identical sized fragments from  $\lambda$ E1 and  $\lambda$ E7, as shown in Figure 2; these are fragments E1.1 and E7.8 of  $\lambda$ E1 and  $\lambda$ E7 genomic clones respectively. Thus the arrangement of genes in the genomic clones is unlikely to be an artefact of the cloning procedure. There are also bands in the genomic DNA of approximately 2.5 kb, 4.8 kb and 8 kb in size which are not found from the digestion of  $\lambda$ E1 or  $\lambda$ E7; these could represent genes such as the 5' sequences of WSBE I-D1 or WSBE I-D3; see below.

Example 3                      Tandem Arrangement of SBE I Type Genes in the *T. tauschii* Genome

Basic restriction endonuclease maps for  $\lambda$ E1 and  $\lambda$ E7 are shown in Figure 3. The map was constructed by performing a series of hybridisations of *Eco*RI or *Bam*HI digested DNA from  $\lambda$ E1 or  $\lambda$ E7. The probes used were the fragments generated from *Bam*HI digestion of the relevant clone. Confirmation of the maps was obtained by PCR analysis, using primers both within the insert and also from the arms of lambda itself. PCR was performed in 10  $\mu$ l volume using reagents supplied by Perkin-Elmer. The primers were used at a concentration of 20  $\mu$ M. The program used was 94°C, 2 min, 1 cycle, then 94°C, 30 sec; 55°C, 30 sec; 72°C, 1min for 36 cycles and then 72°C, 5 min; 25°C, 1 min.

Sequencing was performed on an ABI sequencer using the manufacturer's recommended protocols for both dye primer and dye terminator technologies. Deletions were carried out using the Erase-a-base kit from Promega.

Sequence analysis was carried out using the GCG version 7 package of computer programs (Devereaux et al, 1984).

The PCR products were also used as hybridisation probes. The positioning of the genes was derived from sequencing the ends of the *Bam*HI subclones and also from sequencing PCR products generated from primers based on the



insert and the lambda arms. The results indicate that there is only a single copy of a SBE I type gene within  $\lambda$ E1. However, it is clear that  $\lambda$ E7 resulted from the cloning of a DNA fragment from within a tandem array of the SBE I type genes. Of the three genes in the clone, which are named as wSBE I-D1, wSBE I-D2 and wSBE I-D3); only the central one (wSBE I-D2) is complete.

Example 4 Construction and Screening of cDNA Library

10 A wheat cDNA library was constructed from the cultivar Rosella using pooled RNA from endosperm at 8, 12, 18 and 20 days after anthesis.

The cDNA library was prepared from poly A<sup>+</sup> RNA that was extracted from developing wheat grains (cv. Rosella, a hexaploid soft wheat cultivar) at 8, 12, 15, 18, 21 and 30 days after anthesis. The RNA was pooled and used to synthesise cDNA that was propagated in lambda ZapII (Stratagene).

The library was screened with a genomic fragment from  $\lambda$ E7 encompassing exons 3, 4 and 5 (fragment E7.8 in Figure 3). A number of clones were isolated. Of these an apparently full-length clone appeared to encode a novel type of cDNA for SBE I. This cDNA has been termed SBE I-D2 type cDNA. The putative protein product is compared with the maize SBE I and rice SBE I type sequences in Figure 4. The main difference is that this putative protein product is shorter at the C-terminal end, with an estimated molecular size of approximately 74 kD compared with 85 kDa for rice SBE I (Kawasaki et al, 1993). Note that amino acids corresponding to exon 9 of rice are missing in SBE I-D2 type cDNA, but those corresponding to exon 10 are present. There are no amino acid residues corresponding to exons 11-14 of rice; furthermore, the sequence corresponding to the last 57 amino acids of SBE I-D2 type has no significant homology to the sequence of the rice gene.

We expressed SBE I-D2 type cDNA in *E. coli* in order to examine its function. The cDNA was expressed as a fusion protein with 22 N-terminal residues of  $\beta$ -galactosidase and two threonine residues followed by the SBE I-D2 cDNA sequence either in or out of frame. Although an expected product of about 75 kDa in size was produced from only the in-frame fusion, we could not detect any enzyme activity from crude extracts of *E. coli* protein (data not shown). Furthermore the in-frame construct could not complement an *E. coli* strain with a defined deletion in glycogen branching, although other putative branching enzyme cDNAs have been shown to be functional by this assay (data not shown). It is therefore unclear whether the WSBE I-D2 gene in  $\lambda$ E7 codes for an active enzyme *in vivo*.

#### Example 5            Gene Structure in E7

##### **i. Sequence of WSBE I-D2**

We sequenced 9.2 kb of DNA that contained WSBE I-D2. This corresponds to fragments 7.31, 7.8 and 7.18. Fragment 7.31 was sequenced in its entirety (4.1 kb), but the sequence of about 30 bases about 2 kb upstream of the start of the gene could not be obtained because it was composed entirely of Gs. Elevation of the temperature of sequencing did not overcome this problem. Fragments 7.8 (1 kb) and 7.18 (4 kb) were completely sequenced, and corresponded to 2 kb downstream of the last exon detected for this gene. It was clear that we had isolated a gene which was closely related (approximately 95% sequence identity) to the SBE I-D2 type cDNA referred to above, except that the last 200 bp at the 3' end of the cDNA are not present. The WSBE I-D2 gene includes sequences corresponding to rice exon 11 which are not in the cDNA clone. In addition it does not have exons 9, 12, 13 or 14; these are also absent from the SBE I-D2 type cDNA. The first two exons show lower identity to the corresponding exons from rice (approximately 60%) (Kawasaki et al, 1993) than to the other exons (about 80%). A

diagrammatic exon-intron structure of the wSBE I-D2 gene is indicated in Figure 6. The restriction map was confirmed by sequencing the PCR products that spanned fragments 7.18 and 7.8 and 7.8 and E7.31 (see Figure 3) respectively.

5

**ii. Sequence of wSBE I-D3**

This gene was not sequenced in detail, as the genomic clone did not extend far enough to include the 5' end of the sequence. The sequence is of a SBE-I type. The orientation of the gene is evident from sequencing of the relevant *Bam*HI fragments, and was confirmed by sequence analysis of a PCR product generated using primers from the right arm of lambda and a primer from the middle of the gene. The sequence homology with wSBEI-D2 is about 80% over the regions examined. The 2 kb sequenced corresponded to exons 5 and 6 of the rice gene; these sequences were obtained by sequencing the ends of fragments 7.5, 7.4 and 7.14 respectively, although the sequences from the left end of fragment 7.14 did not show any homology to the rice sequences. The gene does not appear to share the 3' end of SBE I-D2 type cDNA, as a probe from 500 bp at the 3' end of the cDNA (including sequences corresponding to exons 8 and 10 from rice) did not hybridise to fragment 7.14, although it hybridised to fragment 7.18 (data not shown).

**iii. Sequence of wSBE I-D1**

This gene was also not sequenced in detail, as it was clear that the genomic clone did not extend far enough to include the 5' sequences. Limited sequencing suggests that it is also a SBE I type gene. The orientation relative to the left arm of lambda was confirmed by sequencing a PCR product that used a primer from the left arm of lambda and one from the middle of the gene (as above). Its sequence homology with wSBE I-D2 ,D3 and D4 (see below) is about 75% in the region sequenced corresponding to a part of exon 4 of the rice gene.

Starch branching enzymes are members of the  $\alpha$ -amylase protein family, and in a recent survey Svensson (1994) identified eight residues in this family that are invariant, seven in the catalytic site and a glycine in a short turn. Of the seven catalytic residues, four are changed in SBE I-D2 type. However, additional variation in the 'conserved' residues may come to light when more plant cDNAs for branching enzyme I are available for analysis. In addition, although exons 9, 11, 12, 13 and 14 from rice are not present in the SBE I-D2 type cDNA (Figure 5), comparison of the maize and rice SBE I sequences indicate that the 3' region (from amino acid residue 730 of maize) is much more variable than the 5' and central regions (Figure 4). The active sites of rice and maize SBE I sequences, as indicated by Svensson (1994), are encoded by sequences that are in the central portion of the gene. When SBE II sequences from *Arabidopsis* were compared by Fisher et al (1996) they also found variation at the 3' and 5' ends. SBE I-D2 type cDNA may encode a novel type of branching enzyme whose activity is not adequately detected in the current assays for detecting branching enzyme activity; alternatively the cDNA may correspond to an endosperm mRNA that does not produce a functional protein.

Example 6                      Cloning of Specific cDNA Regions of Wheat SBEI Using RT-PCR

The first strand cDNAs were synthesized from 1  $\mu$ g of total RNA, derived from endosperm 12 days after pollination, as described by Sambrook et al (1989), and then used as templates to amplify two specific cDNA regions of wheat SBE I by PCR.

Two pairs of primers were used to obtain the cDNA clones BED1 and BED3 (Table 1). Primers used for cloning of BED3 were the degenerate primer NTS5'

35

5' GGC NAC NGC NGA G/AGA C/TGG 3'

(SEQ ID NO. 1)

in which the 5' end is at position 168 of wSBE I-D4 cDNA, as shown in Table 1), based on the N-terminal sequence of wheat SBE I, and NTS3'

5 5' TAC ATT TCC TTG TCC ATCA 3' (SEQ ID NO. 2)

in which the 5' end is at position 1590 of wSBE I-D4 cDNA, (see Table 1), derived from the conserved regions of the nucleotide sequences of BED5 and the maize  
10 and rice SBE I cDNAs. For clone BED1, the primers used were BEC5'

5' ATC ACG AGA GCT TGC TCA (SEQ ID NO. 3)

15 in which the 5' end is at position 1 of wSBE I-D4 cDNA, (see Table 1); the sequence was based on wSBE I-D4, and BEC3'

20 5' CGG TAC ACA GTT GCG TCA TTT TC 3' (SEQ ID NO. 4)

in which the 5' end is at position 334 of wSBE I-D4 cDNA (see Table 1), and the sequence was based on BED 3.

Table 1

Positions of Sequences Relative to WSBE I-D4 Sequences

Sequence Name	WSBE I-D4 Sequence	WSBE I-D4 cDNA Sequence
Putative initiation of translation	4900	11
N-terminal sequence of SBE I	5550	124
End of translated SBE I sequence	10225	2431
End of WSBE I-D4 cDNA sequence	10461	2687
WSBE I-D45	4870, 5860	1,357
WSBE I-D43	9430, 10435	see below
WSBE-I-D43C	see above	2338,2657
WSBE I-D4R	2kb of sequence approx 600bp 3' to position 10461	not applicable
E 1.1	5680, 6400	380,630
BED 1	not referred to	1,354
BED 2	not referred to	169,418
BED 3	not referred to	151,1601
BED 4	not referred to	867,2372
BED 5	not referred to	867,2687

Example 7                      Identification of a gene encoding the major  
N-terminal of SBE I from the endosperm

We have isolated two classes of SBE I genomic clones from *T. tauschii*. One class contained two genomic clone isolates, and this class has been characterised in some detail (Rahman et al, 1997). The complete gene contained within this class of clones was termed *wSBE I-D2*; there were additional genes at either ends of the clone, and these were designated *wSBE I-D1* and *wSBE I-D3*. The other class contained nine genomic clone isolates. Of these  $\lambda$ E1 was arbitrarily taken as a representative clone, and its restriction map is shown in Figure 3; the SBE I gene contained in this clone was called *wSBE I-D4*. Fragments E1.1 (0.8 kb) and E1.2 (2.1 kb) and fragments E1.7 (4.8 kb) and E1.5 (3 kb) respectively were completely sequenced. Fragment E1.7 was found to encode the N-terminal of the SBE I, which is found in the endosperm as described in Morell et al (1997). This is shown in Figure 4. Using antibodies raised against the N-terminal sequence, Morell et al (1997) found that the D genome isoform was the most highly expressed in the cultivars Rosella and Chinese Spring. We have thus isolated from *T. tauschii* a gene, *wSBE I-D4*, whose homologue in the hexaploid wheat genome encodes the major isoform for SBE I that is found in the wheat endosperm.

All nine genomic clones isolated from *T. tauschii* appear to contain the *wSBE I-D4* gene, or very similar genes, on the basis of PCR amplification and hybridisation experiments. However, the restriction patterns obtained for the clones differ with *Bam*HI and *Eco*RI, among other enzymes, indicating that either the clones represent near-identical but distinct genes or they represent the same gene isolated in distinct products of the *Sau* 3A digest used to generate the library.

35

Example 8                      Investigation of other SBE I genomic clones isolated

All ten members of the E1-like class of SBE I genomic clones were investigated by hybridisation with probes derived from fragment E1.7 (sequence *wsBE I-D45*, encoding the translation start signal and the first 100 amino acids from the N-terminal end and intron sequences; see Table 2) and from fragment E1.5 (sequence *wsBE I-D43*, corresponding largely to the 3' untranslated sequence and containing intron sequences, see Table 1). The results obtained were consistent with one type of gene being isolated in different fragments in the different clones, as shown in Figure 5. The PCR products were obtained from the clones  $\lambda$ E1, 2, 9, 14, 27, 31 and 52. These hybridised to *wsBE I-D45* using primers that amplify near the 5' end of the gene (positions 5590-6162 of *wsBE I-D4*). Sequencing showed no differences in sequence of a 200 bp product.

Analysis of the promoter for *wsBE I-D4* allows us to investigate the presence of motifs previously described for promoters that regulate gene expression in the endosperm. Kreis et al (1985) compared prolamin promoters, and suggested that the presence of a motif approximately -300 bp upstream of the transcription start point, called the endosperm box, was responsible for endosperm-specific expression. The endosperm box was subsequently considered to consist of two different motifs: the endosperm motif (EM) (canonical sequence TGTAAG) and the GCN 4 motif (canonical sequence G/ATGAG/CTCAT). The GCN4 box is considered to regulate expression according to nitrogen availability (Muller and Knudsen, 1993). The *wsBE I-D4* promoter contains a number of imperfect EM-like motifs at approximately -100, -300 and -400 as well as further upstream. However, no GCN4 motifs could be found, which lends support to the idea that this motif regulates response to nitrogen, as starch biosynthesis is not as directly dependent on the nitrogen status of the plant as



storage protein synthesis. Comparison of the promoters for *wSBE I-D4* and *D2* (Rahman et al, 1997) indicates that although there are no extensive sequence homologies there is a region of about 100 bp immediately before the first encoded methionine where the homology is 61% between the two promoters. In particular there is an almost perfect match in the sequence over twenty base pairs CTCGTTGCTTCC/TACTCCACT, (position 4723-4742 of the *wSBE I* sequence), but the significance of this is hard to gauge, as it does not occur in the rice promoter for *SBE I*. The availability of more promoters for starch biosynthetic enzymes may allow firmer conclusions to be drawn. There are putative CAAT and TATA motifs at positions 4870 and 4830 respectively of *wSBE I-D4* sequence. The putative start of translation of the mRNA is at position 4900 of *wSBE I-D4*.

Figure 6 shows the structure of the *wSBE I-D4* gene, compared with the genes from rice and wheat (Kawasaki et al, 1993; Rahman et al, 1997). The rice *SBE I* has 14 exons compared with 13 for *wSBE I-D4* and 10 for *wSBE I-D2*. There is good conservation of exon-intron structure between the three genes, except at the extreme 5' end. In particular the sizes of intron 1 and intron 2 are very different between rice *SBE I* and *wSBE I-D4*.

#### Example 9                      Isolation of cDNA for SBE I

Using the maize starch branching enzyme I cDNA as a probe (Baba et al, 1991), 10 positive plaques were recovered by screening approximately 10<sup>5</sup> plaques from a wheat endosperm cDNA library. On purifying and sequencing these plaques it was clear that even the longest clone (BED5, 1822 bp) did not encode the N-terminal sequence obtained from protein analysis. Degenerate primers based on the protein N-terminal and the sequence from BED5 were then used to amplify the 5' region: this produced a cDNA clone termed BED 3 (Table 1 and Figure 7). This cDNA clone overlapped extensively and had 100% sequence identity with

BED5 and BED4 (Figure 7). As almost the entire protein N-terminal sequence had been included in the primer sequence design, this did not provide independent evidence of the selection of a cDNA sequence in the endosperm that encoded the protein sequence of the main form of SBE I. Using a BED3 to screen a second cDNA library produced BED2, which is shorter than BED3 but confirmed the BED3 sequence at 100% identity between positions 169 and 418 (Figure 7 and Table 1). In addition the entire cDNA sequence for BED3 could be detected at a 100% match in the genomic clone  $\lambda$ E1. Primers based on the putative transcription start point combined with a primer based on the incomplete cDNAs recovered were then used to obtain a PCR product from total endosperm RNA by reverse transcription. This led to the isolation of the cDNA clone, BED1, of 300 bp, whose sequence is shown in Figure 7. By analysing this product, a sequence was again obtained that could be found exactly in the genomic clone  $\lambda$ E1, and which overlapped precisely with BED3.

The N-terminal of the protein matches that of SBE I isolated from wheat endosperm by Morell *et al* (1997), and thus it is likely that *wSBE I-D4* cDNA represents the gene for the predominant SBE I isoform expressed in the endosperm. The encoded protein is 87 kDa; this is similar to proteins encoded by maize (Baba *et al*, 1991) and rice (Nakamura *et al*, 1992) cDNAs for SBE I and is distinct from the *wSBE I-D2* cDNA described previously, in which the encoded protein was 74 kDa (Rahman *et al*, 1997).

Five cDNA clones were sequenced and their sequences were assembled into one contiguous sequence using a GCG program (Devereaux *et al*, 1984). This is illustrated in Figure 8. The intact cDNA sequence, *wSBE I-D4* cDNA, is 2687 bp and contains one large open reading frame (ORF), which starts at nucleotides 11 to 13 and ends at nucleotides 2432 to 2434. It encodes a polypeptide of 807 amino acids with a molecular weight of 87 kDa. Comparison of the amino acid sequence encoded by *wSBE I-D4*

cdNA with that encoded by maize and rice *SBE I* cDNAs showed that there is 75-80% identity between any of two these sequences at the nucleotide level and almost 90% at the amino acid level. Alignment of these three polypeptide sequences, as shown in Figure 9, along with the deduced sequences for pea, potato and *wSBE I-D2* type cDNA, indicated that the sequences in the central region are highly conserved, and sequences at the 5' end (about 80 amino acids) and the 3' end (about 60 amino acids) are variable.

Svensson et al (1994) indicated that there were several invariant residues in sequences of the  $\alpha$ -amylase super-family of proteins to which *SBE I* belongs. In the sequence of maize *SBE I* these are in motifs commencing at amino acid residue positions 341,415,472,537 respectively; these are also encoded in the *wSBE I-D4* sequence (Figure 7), further supporting the view that this gene encodes a functional enzyme. This is in contrast to the results with the *wSBE I-D2* gene, where three of the conserved motifs appear not to be encoded (Rahman et al, 1997).

There is about 90% sequence identity in the deduced amino acid sequence between *wSBE I-D4* cDNA and rice *SBE I* cDNA in the central portion of the molecule (between residues 160 and 740 for the deduced amino acid product from *wSBE I-D4* cDNA). The sequence identity of the deduced amino acid sequence of the *wSBE I-D4* cDNA to the deduced amino acid sequence of *wSBE I-D2* is somewhat lower (85% for the most conserved region, between residues 285 to 390 for the deduced product of *wSBE I-D4* cDNA). Surprisingly, however, *wSBE I-D4* cDNA is missing the sequence that encodes amino acids at positions 30 to 58 in rice *SBE I* (see Figure 8). This corresponds to residues within the transit peptide of rice *SBE I*. A corresponding sequence also occurs in the deduced amino acid sequence from maize *SBE I* (Baba et al, 1991) and *wSBE I-D2* type cDNA (Rahman et al, 1997). Consequently the transit sequence

encoded by *wsBE I-D4* cDNA is unusually short, containing only 38 amino acids, compared with 55-60 amino acids deduced for most starch biosynthetic enzymes in cereals (see for example Ainsworth, 1993; Nair et al, 1997). The  
5 *wsBE I-D4* gene does contain this sequence, but this does not appear to be transcribed into the major species of RNA from this gene, although it can be detected at low relative abundance. This raises the possibility of alternative splicing of the *wsBE I-D4* transcript, and also the question  
10 of the relative efficiency of translation/transport of the two isoforms. The possibility of alternative splicing in both rice and wheat has been considered for soluble starch synthase (Baba et al, Rahman et al, 1995). Alternative splicing of soluble starch synthase would give a transit  
15 sequence of 40 amino acids, which is the same length proposed for the product of *wsBE I-D4* cDNA.

We have previously used probes based on exons 4, 5 and 6 (E7.8 and E1.1, see Rahman et al., 1997) of *wsBE-D2* to probe wheat and *T. tauschii* genomic DNA cleaved with  
20 *Pvu*II and *Bam*HI respectively. This region is highly conserved within rice *SBE I*, *wsBE I-D2* and *wsBE I-D4* and produced ten bands with wheat DNA and five with *T. tauschii* DNA. Neither *Pvu*II nor *Bam*HI cleaved within the probe sequences suggesting that each band represented a single  
25 type of SBE I gene. We have described four SBE I genes from *T. tauschii*: *wsBE I-D1*, *wsBE I-D2*, *wsBE I-D3* and *wsBE I-D4* (Rahman et al, 1997 and this specification), and so we may have accounted for most of the genes in *T. tauschii* and, by extension, the genes from the D genome  
30 of wheat. In wheat, at least two hybridising bands could be assigned to each of chromosomes 7A, 7B and 7D.

Example 10                      Tissue specificity and expression during endosperm development

35                      The 300 bp of 3' untranslated sequence of *wsBE I-D4* cDNA does not show any homology with either the *wsBE I-D2* type cDNA that we have described earlier (Rahman

et al, 1997) or with BE-I from rice, as shown in Figure 7. We have called this sequence *wSBE I-D43C*. It seemed likely that *wSBE I-D43C* would be a specific probe for this class of SBE-I, and thus it was used to investigate the tissue  
5 specificity. The results are shown in Figure 10. A RNA species of about 2700 bases in size was found to hybridise. This is very close to the size of the *wSBE I-D4* cDNA sequence. RNA hybridising to *wSBE-I-D43C* is most abundant at the mid-stage of endosperm development (Figure 10) and  
10 in field grown material is relatively constant during the period 12-18 days, the time at which there is rapid starch and storage protein accumulation (Morell et al, 1995).

The sequence contained within the *wSBE I-D4* gene appears to be expressed only in the endosperm (Figure 9).  
15 We could not detect any expression in the leaf. This could be because another isoform is expressed in the leaf, and/or because the amount of SBE I present in the leaf is much less than what is required in the endosperm. Isolation of SBE I clones from a leaf cDNA library would enable this  
20 question to be resolved.

#### Example 11      Intron-Exon Structure of SBE I

By comparison of the cDNA sequence of SBE I with that of *wSBE I-D4* we can deduce the intron-exon structure  
25 of the gene for the major isoform of SBE I that is found in the endosperm. The structure contains 14 exons compared to 14 for rice (Kawasaki et al, 1993). These 14 exons are spread over 6 kb of sequence, a distance similar to that found in both rice *SBE I* and *wSBE I-D2*. A dotplot  
30 comparison of *wSBE I-D4* sequence and that of rice *SBE I* sequence, depicted in Figure 11, shows good sequence identity over almost the entire gene starting from about position 5100 of *wSBE I-D4*; the identity is poor over the first 5 kb of sequence corresponding largely to the  
35 promoter sequences. The sequence identity over introns (about 60%) is lower than over exons (about 85%).

Example 12      Repeated Sequences in SBE I

Sequencing of *wSBE I-D4* revealed there was a repeated sequence of at least 300 bp contained in a 2kb fragment about 600 bp after the 3' end of the gene. We  
5 have called this sequence *wSBE I-D4R*. This repeated sequence is within fragment E1.5 (Figure 10 and Table 2) and is flanked by non-repetitive sequences from the genomic clone. We have previously shown that the restriction pattern obtained by digesting  $\lambda$ E1 with the restriction  
10 enzyme *Bam*HI is also obtained when *T. tauschii* DNA is digested. Thus *wSBE I-D4R* is unlikely to be a cloning artefact. A search of the GeneBank Database searches revealed that *wSBE I-D4R* shared no significant homology with any sequence in the database. Hybridisation  
15 experiments with *wSBE I-D4R* showed that all of the other *SBE I-D4* type genomic clones (except number 29) contained this repeated sequence (data not shown). The *wSBE I-D4R* sequence was not highly repeated and occurred in the wheat genome with a similar frequency as the *wSBE I-D4* sequence.  
20 When *SBE I-D4R* was used as the probe on wheat DNA from the nulli-tetra lines, four bands were obtained; two of these bands could be assigned to chromosome 7A and the others to chromosomes 7B and 7D (Figure 11C). One of the two *Bam*HI fragments from wheat DNA which could be assigned  
25 to chromosome 7A was distinct from the single band from chromosome 7A detected using *wSBE I-D43* as the probe; the other three bands coincided in the autoradiograph with bands obtained with *wSBE I-D43*, and are likely to represent the same fragment. However, one of these fragments was  
30 distinct from the *Bam*HI fragment that hybridised to the *wSBE I-D43* sequence. In *wSBE I-D4* the *wSBE I-D43* sequence is only 300 bp upstream of *wSBE I-D4R*, and occurs in the same *Bam*HI fragment. These results suggest that the *wSBE I-D4R* sequence can occur independently of *wSBE I-D4* in  
35 the wheat genome.

Example 13                    Isolation of Genomic Clones Encoding SBE II

Screening of a cDNA library prepared from the wheat endosperm with the maize BE I clone (Baba et al, 1991) at low stringency led to the isolation of two classes of positive plaques. One class was strongly hybridising, and led to the isolation of wheat SBE I cDNA clones, as described in Example 5 and in Rahman et al (1997). The second class was weakly hybridising, and one member of this class was purified. This weakly hybridising clone was termed SBE-9, and on sequencing was found to contain a sequence that was distinct from that for SBE I. This sequence showed greatest homology to maize BE II sequences, and was considered to encode part of the wheat SBE II sequence.

The screening of approximately  $5 \times 10^5$  plaques from a genomic library constructed from *T. tauschii* with the SBE-9 sequence led to the isolation of four plaques that were positive. These were designated *wSBE II-D1* to *wSBE II-D4* respectively, and were purified and analysed by restriction mapping. Although they all had different hybridization patterns with SBE-9, as shown in Figure 15, the results were consistent with the isolation of the same gene in different-sized fragments.

Example 14                    Identification of the N-terminal sequence of SBE II

Sequencing of the SBE II gene contained in clone 2, termed *SBE II-D1*, showed that it coded for the N-terminal sequence of the major isoform of SBE II as identified by Morell et al (1997). This is shown in Figure 16a.

Example 15                    Intron-Exon Structure of the SBE II Gene

In addition to encoding the N-terminal sequence of SBE II, as shown in Example 10, the cDNA sequence reported by Nair et al (1997) was also found to have 100% sequence identity with part of the sequence of *wSBE II-D1*.

Thus the intron-exon structure can be deduced, and this is shown in Figure 17.

Example 16                      Number of SBE II Genes in *T. tauschii* and Wheat

5

Hybridisation of the SBE II conserved region with *T. tauschii* DNA revealed the presence of three gene classes. However, in our screening we only recovered one class. Hybridisation to wheat DNA indicated that the locus  
10 for SBE II was on chromosome 2, with approximately 5 loci in wheat; most of these appear to be on chromosome 2D, as shown in Figure 18.

Example 17                      Expression of SBE II

15

Investigation of the pattern of expression of SBE II revealed that the gene was only expressed in the endosperm. However the timing of expression was quite distinct from that of SBE I, as illustrated in Figure 10.

Whereas SBE I gene expression is only clearly  
20 detectable from the mid-stage of endosperm development (Figure 7), SBE II gene expression is clearly seen much earlier, in endosperm tissue at 5-8 days after development (Figure 10), corresponding to an early stage of endosperm development in Figure 7.

25

Example 18                      RT-PCT Amplification of SSS cDNA Sequence from Wheat

A conserved sequence region was used for the synthesis of primers for amplification of SSS by comparison  
30 with the nucleotide sequences encoding soluble starch synthases of rice and pea. A 300 bp RT-PCR product was obtained by amplification of cDNA from wheat endosperm at 12 days post anthesis. The 300 bp RT-PCT product was then cloned, and its sequence analysed. The comparison of its  
35 sequence with rice SSS cDNA showed about 80% sequence homology. The 300 bp RT-PCR product was 100% homologous to



the partial sequence of a wheat SSS in the database produced by Block et al (1997).

Example 19                      Cloning of Wheat Soluble Starch Synthase  
5                                      CDNA

The 300 bp cDNA fragment of wheat soluble starch synthase isolated in Example 14 was used as a probe for the screening of a wheat endosperm cDNA library (Rahman et al, 1996). Eight cDNA clones were selected. One of the  
10 largest cDNA clones was used for DNA sequencing analysis, and gave a 2662 bp nucleotide sequence. A large open reading frame of this cDNA encoded a 647 amino acid polypeptide, starting at nucleotides 247 to 250 and terminating at nucleotides 2198 to 2200. The deduced  
15 polypeptide was shown by protein sequence analysis to contain the N-terminal sequence of a 75 kDa granule-bound protein (Rahman et al, 1995). This is illustrated in Figures 21a and 21b. The location of the 75 kDa protein was determined both the soluble fraction and starch  
20 granule-bound fraction by the method of Denyer et al (1995). Thus this cDNA clone encoded a polypeptide comprising a 41 amino acid transit peptide and a 606 amino acid mature peptide. The cleavage site LRRL was located at amino acids 36 to 39 of the transit peptide of this deduced  
25 polypeptide.

Comparison of wheat SSS with rice SSS and potato SSS showed that there is 87.4% or 75.9% homology at the amino acid level and 74.7% or 58.1% homology at the nucleotide level. Some amino acids in the at N-terminal  
30 sequences of the SSS of wheat and rice were conserved.

Example 20                      Isolation of Genomic Clone of Wheat Soluble  
   Starch Synthase

Seven genomic clones were obtained with a 300 bp  
35 cDNA probe by screening approximately  $5 \times 10^5$  plaques from a genomic DNA library of *Triticum tauschii*, as described above. DNA was purified from 5 of these clones and

digested with *Bam*HI and *Sac*I. Southern hybridization analysis using the 300 bp cDNA as probe showed that these clones could be classified into two classes, as shown in Figure 23. One genomic clone, sg3, contained a long insert,  
5 and was digested with *Bam*HI or *Sac*I and subcloned into pBluescript KS+ vector.

These subclones were analysed by sequencing, and the sequence of the genomic clone sg3 is shown in Figure 21c

10 Example 21      Northern Hybridization Analysis of the  
Expression of Genes Encoding Soluble Starch  
Synthase

Total RNAs were purified from leaves, pre-anthesis material, and various stages of developmental  
15 endosperm at 5-8, 10-15 and 18-22 days post anthesis. Northern hybridization analysis showed that mRNAs encoding wheat SSS were specifically expressed in developmental endosperm. Expression of this mRNAs in the leaves and pre-anthesis materials could not be detected by northern  
20 hybridization analysis under this experimental condition. Wheat SSS mRNAs started to express at high levels at an early stage of endosperm, 5-8 days post anthesis, and the expression level in endosperm at 10-15 days post anthesis, was reduced. These results are summarized in Figure 24.

25 Example 22      Genomic Localisation of Wheat Soluble  
Starch Synthase

DNA from chromosome engineered lines was digested with the restriction enzyme *Bam*HI and blotted onto  
30 supported nitrocellulose membranes. A probe prepared from the 3' end of the cDNA sequence, from positions 2345 to 2548, was used to hybridise to this DNA. The presence of a specific band was shown to be associated with the presence of chromosomes 7A (Figure 25). These data demonstrate  
35 location of the SSS gene on chromosome 7.

Example 23            Isolation of SSS Promoter

We have isolated the promoter that drives this pattern of expression for SSS. The pattern of expression for SSS is very similar to that for SBE II: the SSS gene transcript is detectable from an early stage of endosperm development until the endosperm matures. The sequence of this promoter is given in Figure 26.

Example 24            Isolation of the Gene Encoding Debranching Enzyme from Wheat

The *sugary* mutation in maize results in mature dried kernels that have a glassy and translucent appearance; immature mature kernels accumulate sucrose and other simple sugars, as well as the water-soluble polysaccharide phytoglycogen (Black et al, 1966). Most data indicates that in *sugary* mutants the concentration of amylose is increased relative to that of amylopectin. Analysis of a particular *sugary* mutation (*su-Ref*) by James et al, (1995) led to the isolation of a cDNA that shared significant sequence identity with bacterial enzymes that hydrolyse the  $\alpha$  1,6-glucosyl linkages of starch, such as an isoamylase from *Pseudomonas* (Amemura et al, 1988), *ie.* bacterial debranching enzymes.

We have now isolated a sequence amplified from wheat endosperm cDNA using the polymerase chain reaction (PCR). This sequence is highly homologous to the sequence for the *sugary* gene isolated by James et al, (1995). This sequence has been used to isolate homologous cDNA sequences from a wheat endosperm library and genomic sequences from *Triticum tauschii*.

Comparison of the deduced amino acid sequences of DBE from maize with spinach (Accession SOPULSO, unpublished), *Pseudomonas* (Amemura et al, 1988) and rice (Nakamura et al, 1997) enabled us to deduce sequences which could be useful in wheat. When these sequences were used as PCR amplification primers with wheat genomic DNA a product of 256 bp was produced. This was sequenced and was

compared to the sequence of maize *sugary* isolated by James et al, (1995). The results are shown in Figure 28a and 28b. This sequence has been termed wheat debranching enzyme sequence I (WDBE-I).

5           WDBE-1 was used to investigate a cDNA library constructed from wheat endosperm (Rahman et al, 1997) enables us to isolate two cDNA clones which hybridise strongly to the WDBE-I probe. Use of WDBE 1 to investigate a genomic library constructed from *T. tauschii*, as  
10 described above has led to the isolation of four genomic clones which hybridised strongly to the WDBE-I sequence. Hybridization of WDBE-I to DNA from *T. tauschii* indicates one hybridizing fragment (Figure 29).

          We have clearly isolated a sequence from the  
15 wheat genome that has high identity to the debranching enzyme cDNA of maize characterised by James et al (1997). The isolation of homologous cDNA sequences and genomic sequences enables further characterisation of the debranching enzyme cDNA and promoter sequences from wheat  
20 and *T. tauschii*. These sequences and the WDBE I sequences shown herein are useful in the manipulation of wheat starch structure through genetic manipulation and in the screening for mutants at the equivalent *sugary* locus in wheat.

25   Example 25           Use of probes from granule-bound starch synthase and SBE II sequences to identify null or altered alleles for use in breeding programmes

          There are two general strategies for obtaining  
30 wheats with altered starch structure:

- (a) using genetic engineering strategies to suppress the activity of a specific gene, or to introduce a novel gene into a wheat line.
- (b) selecting among existing variation in wheat  
35 for missing ("null") or altered alleles of a gene in each of the genomes of wheat, and combining these by plant breeding.

DNA primer sets were designed to enable amplification of the first 9 introns of the SBE II gene using PCR. The design of the primer sets is illustrated in Figure 30. Primers were based on the wSBE II-D1 sequence (Figure 19) and were designed such that intron sequences in the wSBE II sequence (deduced from Figure 17) were amplified by PCR. These primer sets individually amplify the first 9 introns of SBE II. One primer set, for intron 6, was found to amplify products from each of chromosomes 2A, 2B and 2D of wheat. This is shown in Figure 31, which illustrates results obtained with chromosome 2 nullisomic tetrasomic lines of the cultivar Chinese Spring.

Figure 32 compares results of amplification with the Intron 6 primer set for normal lines of the cultivars Chinese Spring and Rosella. In Chinese Spring a PCR product of 213 bp is absent, indicating that this cultivar possesses a potential null allele. Thus Chinese Spring can be used as a parental line for breeding programmes for generation of new lines in which expression of SBE II is diminished or abolished, with consequent increase in amylose content of the wheat grain. Thus a high amylose wheat can be produced.

It will be apparent to the person skilled in the art that while the invention has been described in some detail for the purposes of clarity and understanding, various modifications and alterations to the embodiments and methods described herein may be made without departing from the scope of the inventive concept disclosed in this specification.

Reference cited herein are listed on the following pages, and are incorporated herein by this reference.

REFERENCES

- Ainsworth, C., Clark, J. and Balsdon, J.  
Plant Molecular Biology, 1993 22 67-82
- 5 Amemura, A., Chakrabort, R., Fujita, M., Noumi, T. and  
Futai, M.  
J. Biol. Chem., 1988 263 9271-9275
- 10 Baba, T., Kimura, K., Mizuno, K., Etoh, H., Ishida, Y.,  
Shida, O. and Arai, Y.  
Biochem. Biophys. Res. Commun., 1991 181 87-94.
- Black, R.C., Loerch, J.D., McARDle, F.J. and Creech, R.G.  
15 Genetics, 1966 53 661-668
- Burton, R.A., Bewley, J.D., Smith, A.M.,  
Bhattacharya, M.K., Tatge, H., Ring, S., Bull, V.,  
Hamilton, W.D.O. and Martin, C.  
20 The Plant Journal, 1995 7 3-15.
- Cangiano, G., La Volpe, A., Poulsen, P. and Kreiberg, J.D.  
Plant Physiology, 1993 102 1053-1054.
- 25 Clarke, B.C., Mukai, Y. and Appels, R.  
Chromosoma, 1996 105 269-275
- Devereaux, J., Haeberli, P. and Smithies, O.  
Nucleic Acids Res., 1984 12, 387-395.
- 30 Denyer, K., Hylton, C.M., Jenner, C.F. and Smith, A.M.  
Planta, 1995 196 256-265
- Doherty, J.P., Lindeman, R., Trent, R.J., Graham, M.W. and  
35 Woodcock, D.M.  
Gene, 1992 124 113-120

- Gill, B.S. and Appels, R.  
Plant Syst. Evol., 1988 160 77-90.
- Jahne, A., Lazzeri, P.A., Jager-Gussen, M. and Lorz, H.  
5 Theor. Appl. Genet., 1991 82 47-80
- James, M.G., Robertson, D.S. and Myers, A.M.  
Plant Cell, 1995 7 417-429
- 10 Jolly, C.J., Glenn, G.M. and Rahman, S.  
Proc. Natl Acad. Sci., 1996 93 2408-2413.
- Kawasaki, T., Mizuno, K., Baba, T. and Shimada, H.  
Molec. Gen. Genet., 1993 237 10-16.
- 15 Lagudah, E.S., Appels, R. and McNeill, D.  
Genome, 1991 34 387-395
- Lazzeri, P.A., Brettschneider, R., Luhrs, R. and Lorz, H.  
20 Theor. Appl. Genet., 1991 81 437-444
- Maniatis, T., Fritsch, E.F. and Sambrook, J.  
Molecular cloning. A Laboratory Manual., New York. Cold  
Spring Harbor Laboratory, 1982
- 25 Mizuno, K., Kawasaki, T., Shimada, H., Satoh, H.,  
Koyabashi, E., Okumura, S., Arai, Y. and Baba, T.  
J.Biol. Chem., 1993 268 19084-19091.
- 30 Martin, C. and Smith, A.  
The Plant Cell, 1995 7 971-985.
- Morell, M.K., Blennow, A., Kosar-Hashemi, B. and  
Samuel, M.S.  
35 Plant Physiol., 1996 113 201-208.

- Morell, M.K., Rahman, S., Abrahams, S.L. and Appels, R.  
Aust.J.of Plant Physiol., 1995 22 647-660.
- 5 Nair, R., Baga, M., Scoles, G.J., Kartha, K. and  
Chibbar, R.  
Plant Science, 1997 1222 153-163
- 10 Nakamura, Y., Takeichi, T., Kawaguchi, K. and  
Yamanouchi, H.  
Physiologia Plantarum, 1992 84 329-335.
- Nakamura, Y., Umemoto, T. and Sasaki, T.  
Planta, 1996 199 209-214
- 15 Preiss, J.  
Biology and Molecular Biology of starch synthesis and its  
regulation. In 'Oxford Surveys of Plant Molecular and Cell  
Biology., 1991 Vol. 7.' (Ed. B. J. Mifflin.) pp. 59-  
114. (Oxford University Press: Oxford.)
- 20 Rahman, S., Kosar-Hashemi, B., Samuel, M., Hill, A.,  
Abbott, D.C., Skerritt, J.H., Preiss, J., Appels, R. and  
Morell, M.  
Aust. J. Plant Physiol., 1995 22 793-803.
- 25 Rahman, S., Abrahams, S., Mukai, Y., Abbott, D.,  
Samuel, M., Morell, M. and Appels, R.  
Genome, 1997 40 465-474
- 30 Sambrook, J., Fritsch, E.F. and Maniatis, T.  
Molecular Cloning: A Laboratory Manual (Cold Spring Harbor  
Laboratory Press, 2<sup>nd</sup> ed 1989)
- Svensson, B.
- 35 Plant Mol. Biol., 1994 25 141-157.



Tingay, S., McElroy, D., Kalla, R., Fieg, S., Wang, M.,  
Thornton, S. and Bretell, R.

The Plant Journal, 1997 11 1369-1376

5 Wan, Y. and Lemaux, P.G.

Plant Physiology, 1994 104 37-48

COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH  
ORGANISATION and

10 THE AUSTRALIAN NATIONAL UNIVERSITY

12 September 1997

1/66

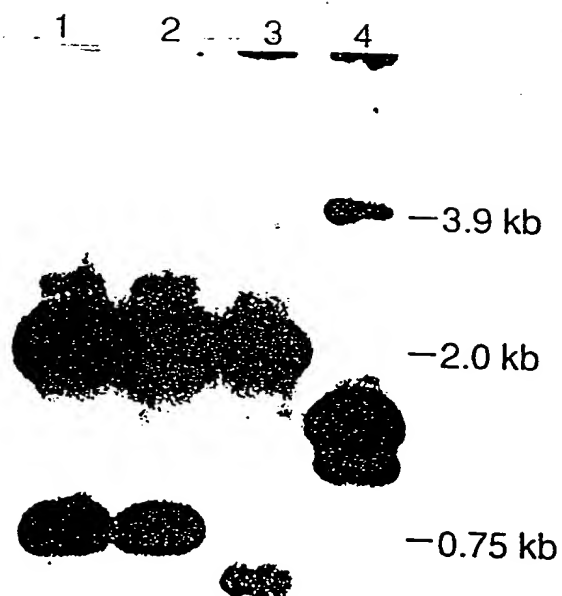


FIGURE 1

2/66

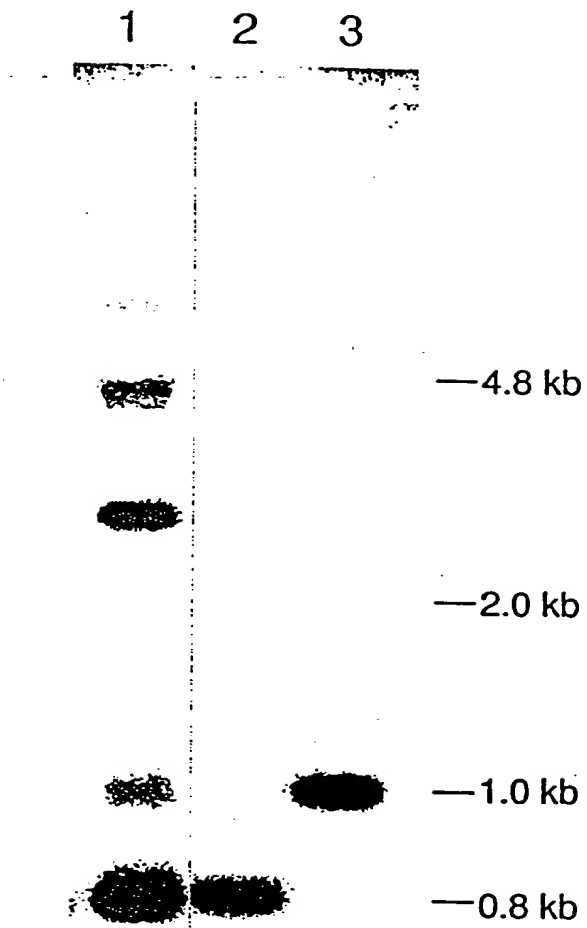


FIGURE 2

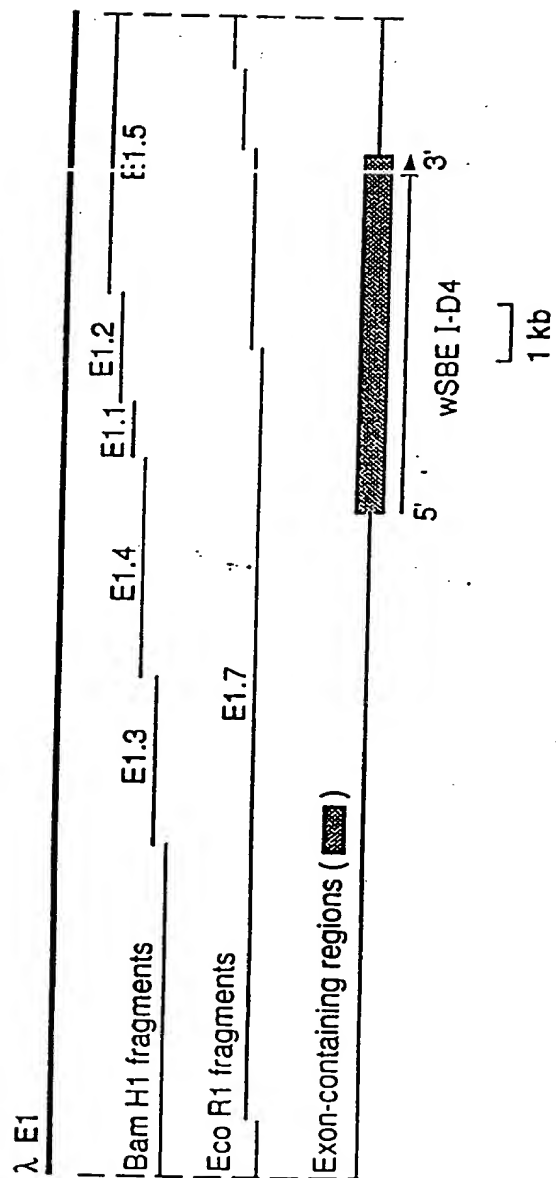
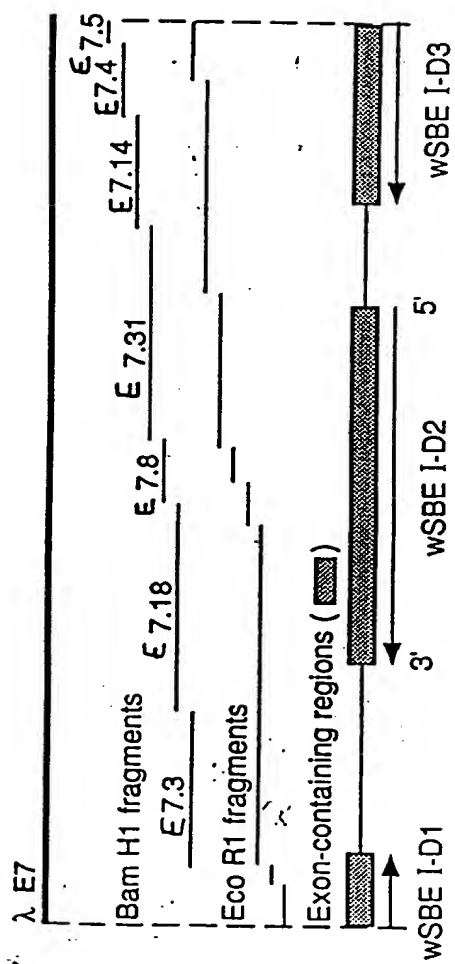


FIGURE 3

5' TCCCGTGTCTGCGCCAAGAGACTACACCATGGCAACAGCTGAAGATGGTGTGGCGACCT 5'  
 3' AGGGCACAGACGGGTTCTCTGATGTGGTACCGTTGTGCGACTTCTACCAACCGCTGGA 3'

DNA

[ S R V C A K R L H H G N S \* R W C W R P  
 P V S A P R D Y T M A T A E D G V G D L  
 P C L R Q E T T P W Q Q L K M V L A T F ]

possible  
 reading  
 frames

[ V S A P R D Y T M A T A E D G V  
 true N-  
 terminal  
 sequence  
 for BE-1  
 (Morell et  
 al, 1997) ]

FIGURE 4

5/66

A

1 2 3 4 5 6 7 8 9 10 11 12 13



B

1 2 3 4 5 6 7 8 9 10 11 12

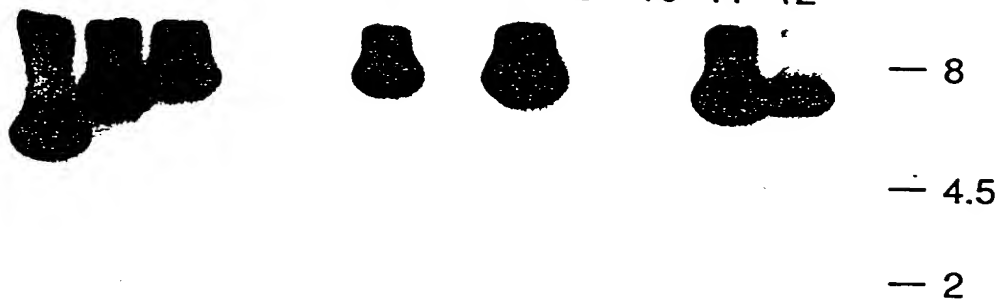


FIGURE 5

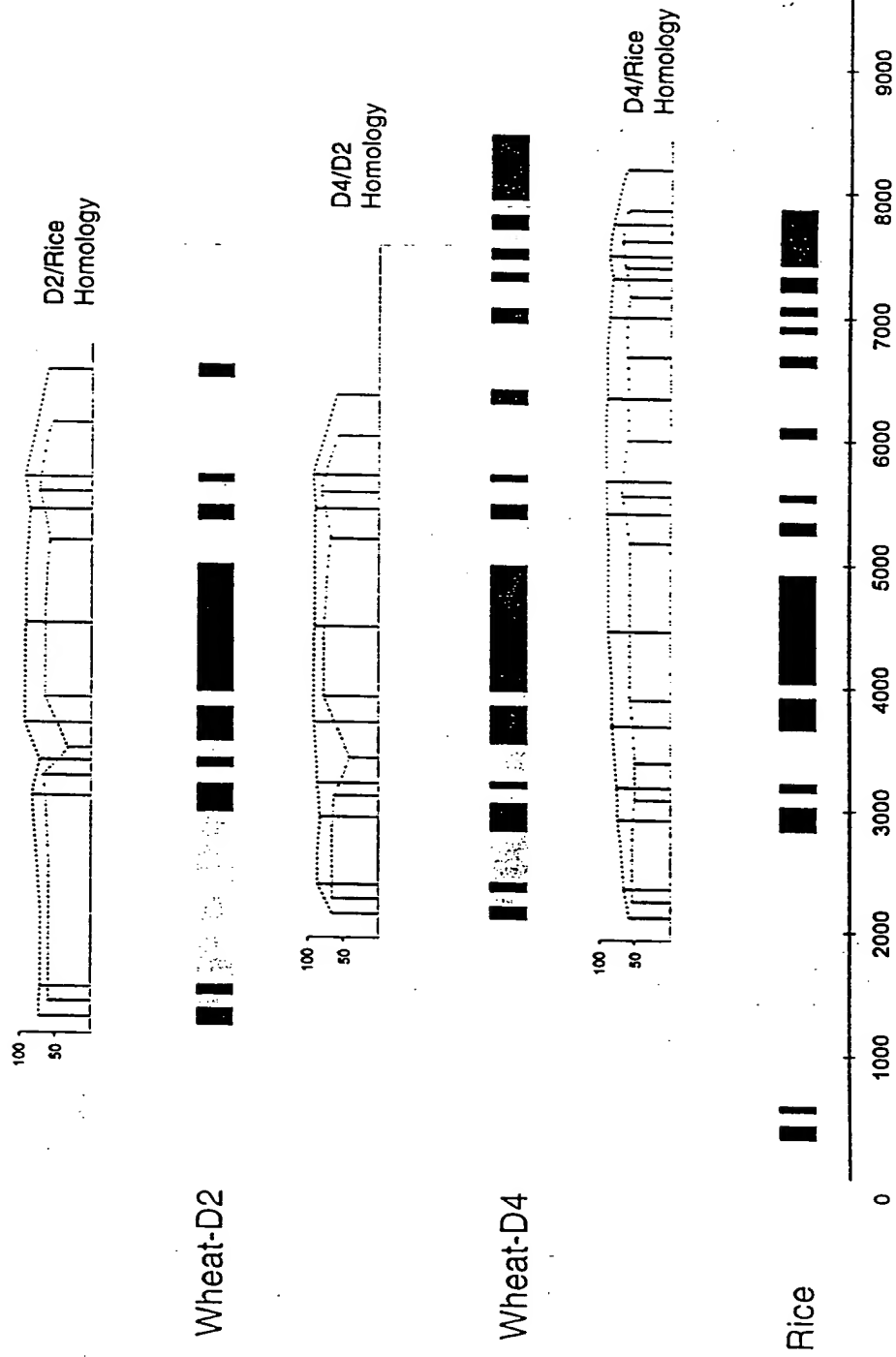


FIGURE 6

7/66

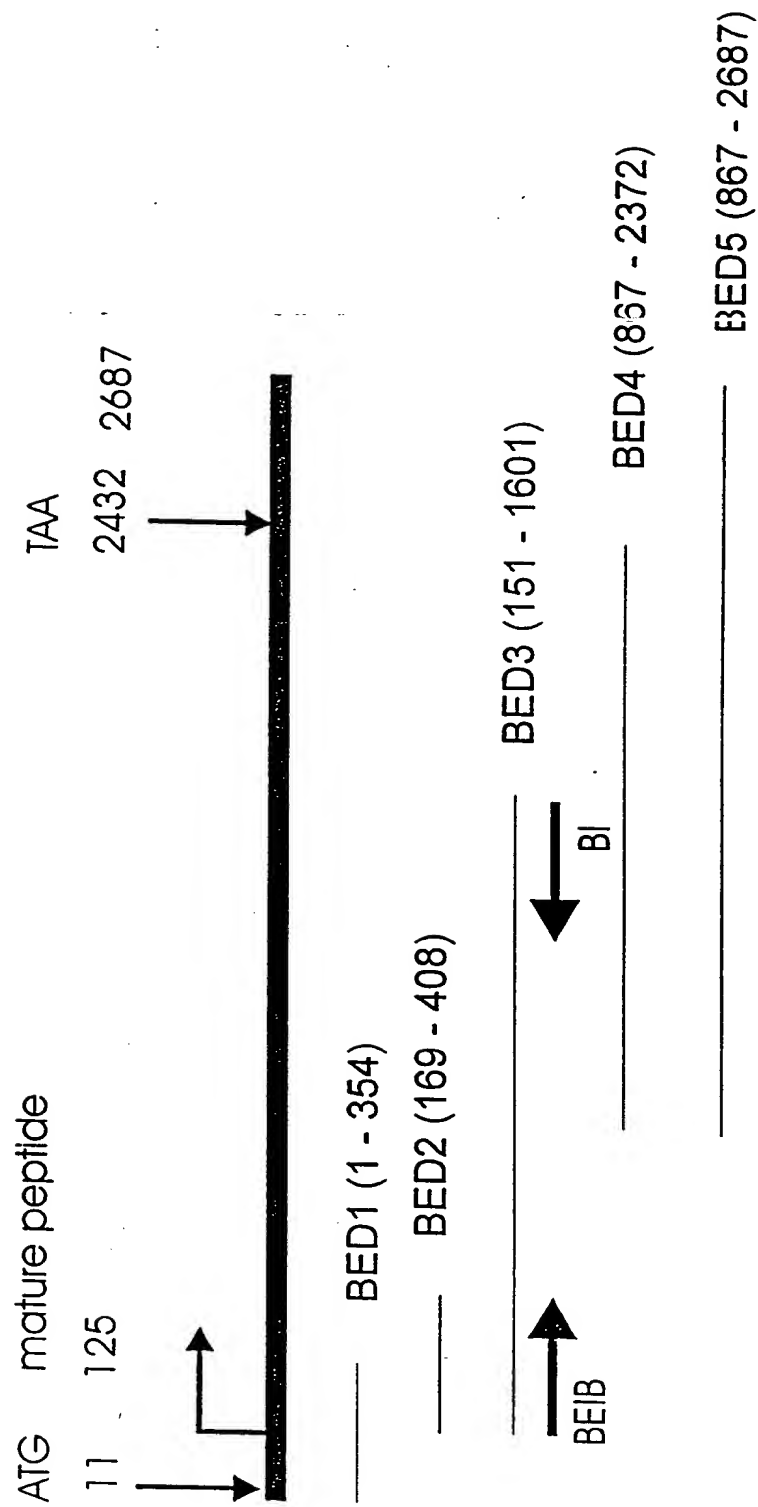


FIGURE 7



8/66

1 MLCLTAPSCS PSLPPRPSRP AADRPGGIS AKSKFSVPVS APRDYMATA  
51 EDGVGDLPIY DLDPKFAGFK EHFSYRMKKY LDQKHSIEKH EGGLEEFSKG  
101 YLKFGINTEN DATVYREWAP AAMDAQLIGD FNNWNGSGHR MTKDNYGVWS  
151 IRISHVNGKP AIPHNSKVKF RFHRGDGLWV DRVPAWIRYA TFDASKFGAP  
201 YDGVHWDPPS GERYVFKHPR PRKPDAPRIY EAHVGMSGER PEVSTYREFA  
251 DNVLPRIKAN NYNTVQLMAI MEHSILCFFW YHVTNFFAVS SRSGTPEDLK  
301 YLVDKAHS LG LRVLMQVHS HASSNMTDGL NGYDVGQNTQ ESYFHTGERG  
351 YHKLWDSRLF NYANWEVLRY LLSNLRYWMD EFMFDGFRFD GVTSMLYNHH  
401 GINMSFAGNY KEYFGLD TDV DAVVYMMLAN HLMHKILPEA TVVAEDVSGM  
451 PVL CRSVDEG GVGFDYRLAM AIPDRWIDYL KNKDDLEWSM SAIAHTLTNR  
501 RYTEKCIAYA ESHDQSIVGD KTMAFLIMDK EMYTGMSDLQ PASPTIDRGI  
551 ALQKMIHFIT MALGGDGYLN FMGNEFGHPE WIDFPREGNN WSYDKCRRQW  
601 SLSDIDHLRY KYMNAFDQAM NALDDKFSFL SSSKQIVSDM NEEKKIIVFE  
651 RGD LVFVFNF HPSKTYDGYK VGCDLPGKYK VALDSDALMF GGHGRVAQYN  
701 DHFTSPEGVP GVPETNFNMR PNSFKVLSPP RTCVAYYRVE EKAENLRMKE  
751 LLLGAKLLLG TSMLKPLVSK TQOMVRLLV PKRRLOGGDS SKKGINFVFG  
801 SPDKDNK\*

FIGURE 8

9/66

	1				50
RSBEI	.....	*****	*...***pl	lp*****	**ag*****
MSBEI	.....	*****v*p**	**tplp***r	***h***aa*	pg*****
D4cDNA	.....	*****ap*c	**sl...***p	**pa*****g*	**s*.....
PESBEII	.....	.....	.....	.....	.....
POSBE	meinfkvlsk	pirgsfp*f*	pkv*sgas*n	kic*psql*t	*lkf*sqers
D2cDNA	.....	*****s*ll	prp*a*....	.....***l*	*****ggk
Consensus	-----	MLCLTSGGS	SP-S-APPR-	SRS-ADRPSP	GIIAGGGNVR
	51				100
RSBEI	1..**v*...	*p*****g**	*tn***pa**	rk*****v*v	***.*****
MSBEI	1..**l**qc	ka***gv***	****ataa*v	q*d*****ak	g***.*****
D4cDNA	.....	.....	*****p*s*	prdy*****a*	*g*..gd***
PESBEII	.....	.....	.....mt	d*ks**psv*	**f*..nig*
POSBE	w..d*s*t*k	*rv*kde*mk	h*saisa*lt	d**s***pl*	***kt*nigl
D2cDNA	rlsv*p***f	ll**l****a	***sf*s***	rg**ia**..	tgys*****
Consensus	---SV-SVP-	S-RRSWPRKV	KSKFSV-VTA	-DNKTMAT-E	EDV--DHLPI
	101				150
RSBEI	*****e*	****n**i**	*****c****	*****v	*****v
MSBEI	*****i*	*****g**e	n**s**s***	*****n	*****n
D4cDNA	*****ag*	****s****k	*****g***	*****s	*****s
PESBEII	lnv**ss**p*	****k*****	**h**k****e	y****q**a*	*****f*r*
POSBE	ln****t**p*	l****h****	*v***m****	y**p****aq	*****f*r*
D2cDNA	****l**ae*	****d*trn*	*i*****g	*****g	*****g
Consensus	YDLDPKLE-F	KDHFYRMRK	YLDQKHLIEK	HEGGLEEFK	GYLKFGINTE
	151				200
RSBEI	*g*****	*****ak*	*****k****	**k*****	*****k
MSBEI	*dg*****	*****e***	***d***a**	*****k****	**k*d**k**
D4cDNA	nd*****	*****m*****	*****g*	r*t**n****	*****n
PESBEII	*dgis*****	*****i**	***g*****l	h****q****	**q*pdad*n
POSBE	*gci*****	*****dev**	***g*****	m****q****	*****pd*ds*
D2cDNA	hg*s*****	*****e*****	*****g*	**a**n****	*****n
Consensus	--ATVYREWA	PAAQEAQLIG	DFNNWNGSNH	KMEKD-FGVW	SIRISHVNGK
	201				250
RSBEI	*****r**g*a*	*****f*****	*****f*****	*****f*****	*****f*****
MSBEI	*****l*.g***	*****l***	*****l***	*****l***	*****l***
D4cDNA	*****hr*d*l*	*****f*****	*****f*****	*****f*****	*****f*****
PESBEII	*****r**k*sd***	*****k*	****ptr*a*	*****y****	*****y****
POSBE	*v*****r**k*n***	*****k*	**a**t**a*	*****y****	*****y****
D2cDNA	*****r*.h***	**q*****	****t**es**	*****l*****	*****l*****
Consensus	PAIPHNSKVK	FRF-HG-GVW	VDRIPAWIRY	ATVDASKFGA	PYDGVHWDPP
	251				300
RSBEI	ac*****	*****k*a*****	*****k*a*****	*****k*a*****	*****k*a*****
MSBEI	a****t****s**a****	*****r*****	*****r*****	*****r*****	*****r*****
D4cDNA	sg*****	*****ss	***r*ns****	**d*****e	**d*****e
PESBEII	l****q****k****	*****ss	***r*ns****	**d*****k*	**d*****k*
POSBE	p****h**y****r****	*****ss	***r*ns****	**d*****k*	**d*****k*
D2cDNA	s*****n**v****	*****v**g	kl*ag*****	p*****cl**	p*****cl**
Consensus	-SERYVFKHP	RPPKPDAPRI	YEAHVGMGE	EPEVSTYREF	ADNVLPPIRA

FIGURE 9

10/66

	301				350
RSBEI	*****	*****	*****	*****	*****
MSBEI	*****	*****	*****	*****	*****
D4cDNA	*****	*****ilcf*	w*****	*****	*****
PESBEII	*****	*****	w*****kp**	*****g**	*****
POSBE	*****	*****g**	*****	*****y*n**	*****
D2cDNA	t*****g	*****ds**	*****	*****	*****
Consensus	NNYNTVQLMA	IMEHSYYASF	GYHVTN-FFA	VSSRSGTPED	LKYL-DKAHS
	351				400
RSBEI	*****	*****	*****u	*li*****t**	*****
MSBEI	*****	*****	*****	*****a**	*****
D4cDNA	*****	*****s*m**	*****n	*****t**	*****
PESBEII	***n*****	*****	*****	s*q*****a**	*****
POSBE	***j**v***	*****	*****	*****	*****
D2cDNA	*****	*****i*	*****	*****	*****
Consensus	LGLRVLMDVV	HSHASNNTVD	GLNGYDVGQS	TQESYFH-GD	RGYHKLWDSR
	401				450
RSBEI	*****	*****	*****	*****	*****k****
MSBEI	*****	*****	*****	*****	*****v****
D4cDNA	*****	*****	*****	*****n	*****s*a*
PESBEII	*****ks.	s*****	****k*****	*****	*****a****
POSBE	*****	*****	****n*****	*****v	*****
D2cDNA	*****	*****	*****	*v*****n	*n*****s*n*
Consensus	LFNYANWEVL	RFLLSNLRYW	-DEFMFDGFR	FDGVTSMLYH	HHGINMGFTG
	451				500
RSBEI	*****	*****	*****l**	*****	*****
MSBEI	**q*****	a*****	*****l**	*****	*****
D4cDNA	*****g***	*****	*****i**	*****	*****s**
PESBEII	d*n*****e**	*****	**s*v*di**	***d*****	***g*g***s
POSBE	**n*****ea*	*****	**n*i*i**	*****	***g*g***s
D2cDNA	*****ig***	n***f*****	*****l**	**i***v***	*****
Consensus	NYKEYFSLDT	DVDAVVYML	ANHLMHK-LP	EATVVAEDVS	GMPVLCRPVD
	501				550
RSBEI	*****	*****	*****rk*	****.*vq**	*****
MSBEI	*****	*****	*****	**g*.*ah**	*****
D4cDNA	*****	*****	*****l**	***a.*ah**	*****
PESBEII	*v*****	*****k***	*****k***	**k*.*sln*	*****
POSBE	*****	*****k***	*****ne**	**k*.*tss*	*****
D2cDNA	***l*****q	**t*****	**e*g*qq*	***sv*sq**	*****p*f*
Consensus	EGGVGFDYRL	AMAIPDRWID	YLKNKDDSEW	SMSE-I--TL	TNRRYTEKCI
	551				600
RSBEI	*****	*****	*****t***	*****n	*****
MSBEI	*****	*****	*****t***	*****	*****
D4cDNA	*****	*****m****	*****t***	*****	*****
PESBEII	s*****	*****	**e***ss*	c*tml*****	***s*h****
POSBE	*****	*****	*****s***	c*td***v**	*****h****
D2cDNA	***rqn***	**s***m****	**w*t*s***	a*d*d*****	*a*****
Consensus	AYAESHQSI	VGDKTIAFL	MDKEMY-GMS	DLQPASPTID	RGIALQKMIH

FIGURE 9 (cont.)



	601				650
RSBEI	*****	*****	*****	*****	*****
MSBEI	*****	*****	*****	*****	*****
D4cDNA	*****	*****	*****	*****	*****
PESBEII	*****	*****	*****	*****	*****s*i
POSBE	*f*****	*****	*****	**g*****	lt**n****n
D2cDNA	*****s	**k*****	*****	*****	****n*a*s*
Consensus	FITMALGGDG	YLNFMGNEFG	HPEWIDFPRE	GNNWSYDKCR	-RQWSLVDTD
	651				700
RSBEI	*****	*****e	*****	*****k***	*****
MSBEI	*****	*****r	*****	*****	*****
D4cDNA	*****	*****	*****	*****k**	*****
PESBEII	*****	*r***l****	**i*a*t***	**st*n****	*****
POSBE	*****	*r***s****	****a*g***	**s*d*n**	*****
D2cDNA	*****	v**vdtps**	c*****nt	a*h*****g	sa*tk*....
Consensus	HLRYKYMNAF	DQAMNALD-K	FSFLSSSKOI	VSDMNEE-KV	IVFERGDLVF
	701				750
RSBEI	*****n***	k*****	*****	**v*****	*****
MSBEI	*****k***	*****	*****	**v*****	*****
D4cDNA	*****s***	*****	***k*****	**m*****	aqyn*****
PESBEII	*****en**	*****	*****	*te*****	***a*q****
POSBE	*****ku**	*****	*****	*wc*****t	*****
D2cDNA	..*thlrsgc*	*p.....s**	stssc**....	..*gpsnqspf	skpfig*pgc
Consensus	VFNFHP-KTY	EGYKVGCDLP	GKYRVALDSD	AL-FGGHGRV	GHDVDHFTSP
	751				800
RSBEI	**m*****	*****	*****	*****	*****
MSBEI	*****	*****	*****	*****	*****
D4cDNA	*****	*****	*****	*****	*****
PESBEII	*****	*****	*****	*****	****h***v*
POSBE	*****	**g*qipskc	cllrehvwli	telmnacq*1	kitrq*f*vs
D2cDNA	ifcc*lfkge	*.....	*****	*****	*****
Consensus	EG-PGVPETN	FNNRP----	-----	-----NSFKV	LSPPRTCWAY
	801				850
RSBEI	*...****dr	****rggava	s**i.vtey.	.....**e**t	sgetisggwk
MSBEI	*...****ag	agr**hakae	t***sp*es.	.....**k*s	ra....*ske
D4cDNA	*...****ka	*n***ke*1	ga*1*lgtsm	lkp**sk*qq	mvrllvpkr
PESBEII	*...****q	**snnpn*gs	*ee**a*adt	dvar*pdvs*	e*..ed*nld
POSBE	*yqqp*sr*v	trnlkiry*q	*sv**tnacq	klk*trq*f*	v*yyqqpilr
D2cDNA	*****	*****	*****	*****	*****
Consensus	Y---RVDER-	EELR---LL-	-GKTL-A---	----IDVTA-	-S----S---
	851		876		
RSBEI	gs*kd*cg**	*mk****r***	e*c*d.		
MSBEI	dk*atagg**	*wk*arqp**	q*t**.		
D4cDNA	*lqgg*ss**	*in***g*p*	k*n**.		
PESBEII	***dnseav	dagilkver*	vvgn*		
POSBE	**trklkds1	stnist*...	*****		
D2cDNA	*****	*****	*****		
Consensus	R-E--D--KK	G--FVF-SSD	-D-K--		

FIGURE 9 (cont.)

# Expression of starch biosynthetic genes

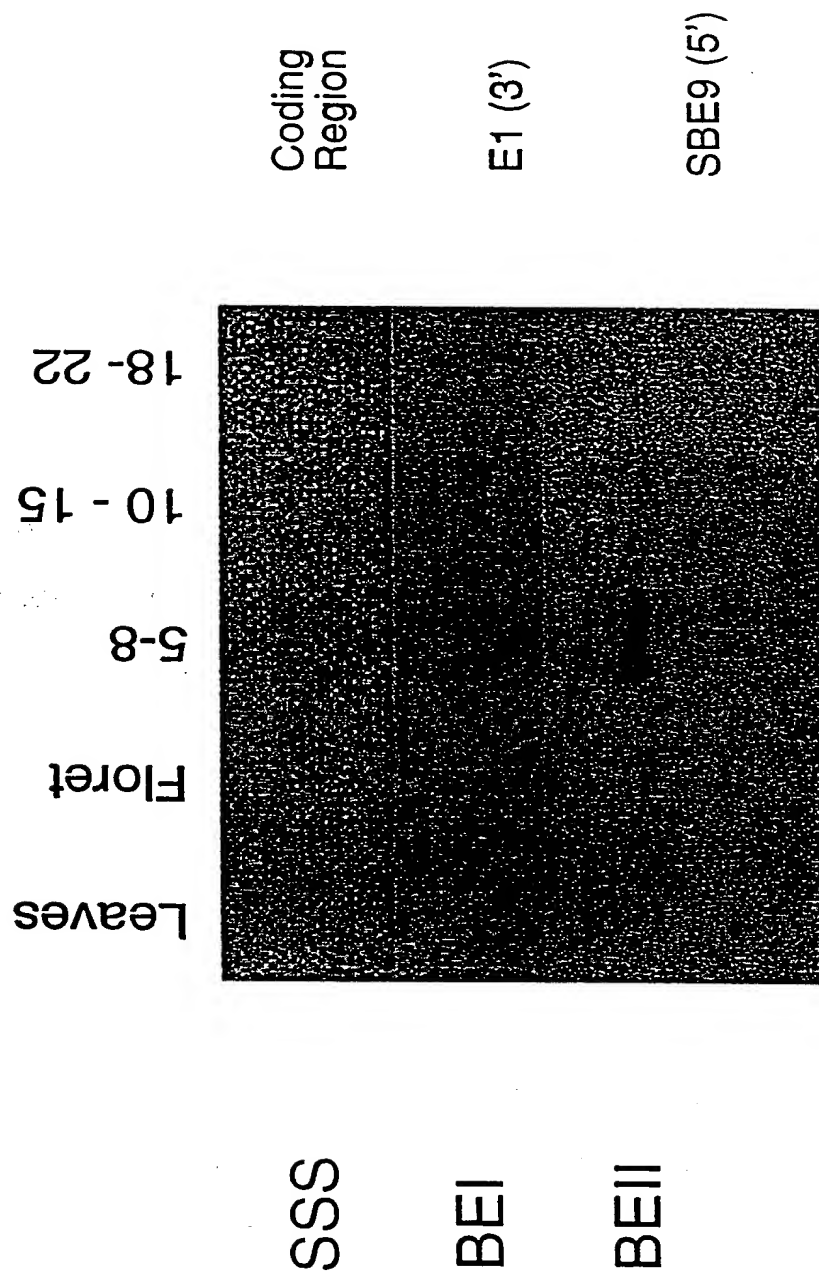


FIGURE 10

13/66

DOTPLOT of: d10838.pnt Density: 12614.77 February 18, 1997 11:43

COMPARE Window: 21 Stringency: 14.0 Points: 20,788

sr427.res ck: 6,362, 1 to 11,099

d10838.em\_pl ck: 3,071, 1 to 11,700

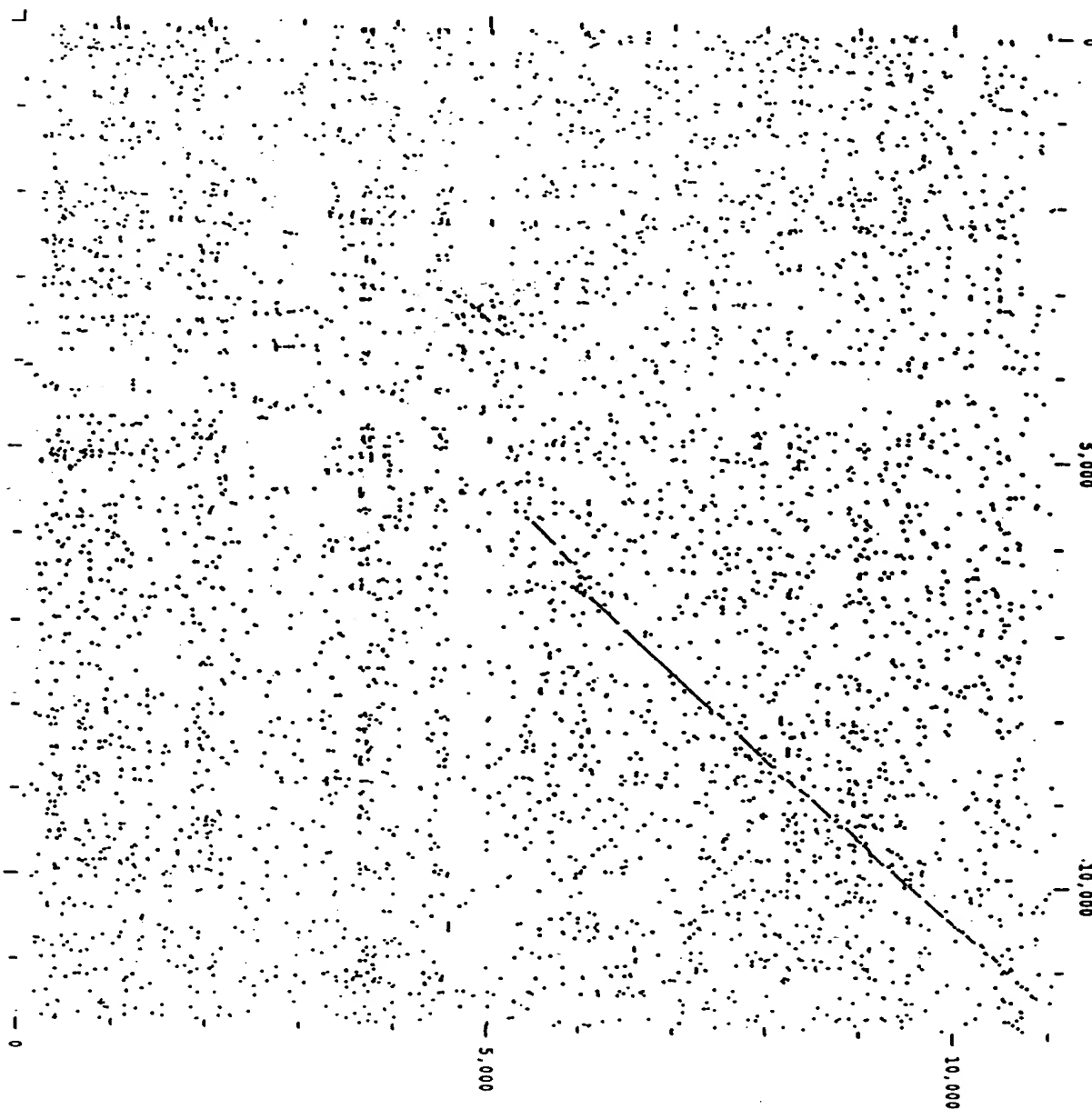


FIGURE 11

14/66

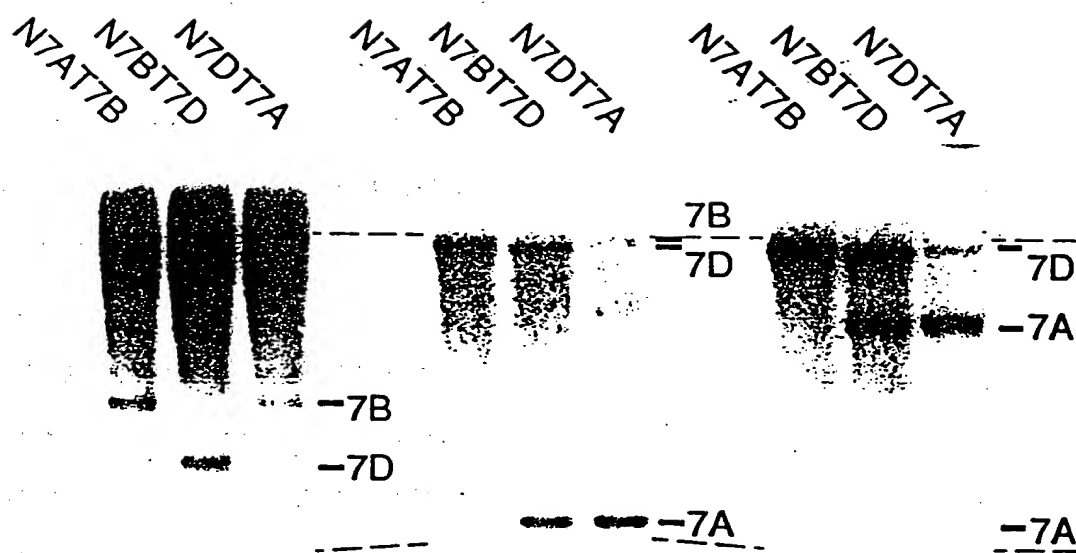


FIGURE 12



15/66

1 GGGTGGCGGG TCGGGCGGCA AGGCGCGGGG CGGCGGGGCG GCNCGGGGCG  
51 GCNCGGGGCG CGCGGGCGGC AGCGGCGGCT AGGGTTTCGC GGCGGCGGCG  
101 ACTTGGGCTG AGGCGGGGCA CGGGCTGCGG CTTTAAAGGC CGGCCAGGCT  
151 GAGGTGTCCG GGTCGGACAC GGCCCGTAAG GCGGTTGACT TTAAAAAATA  
201 ATAATTCGGA CATGCAAAAA AGTAAGAAAA GAAATAATAA ACGGACTCCA  
251 AAAATCCCGA AGTAAATTTT TCCCCATTCT TAAAAATAAG CCGGACAAGA  
301 TGAACATTTA TTTGGGCCTA AAATGCAATT TTGAAAAATG CGTATTTTTT  
351 CTAATTCGGA ATAAAATCAA ATAAAATCCA AATAAAATCA AATATTTGTT  
401 TTTAATATTT TTCTCCAAT ATTTCAATTAT TTGTGAAGAA GTCATTTTAT  
451 CCCATCTCAT ATATTTTGAT ATGAAATATT TTCGGAGAGA AAAATAATTA  
501 AAACAAATGA TCCTATTTTC AAAATTTGAG AAAACCCAAA TATGAAAATA  
551 ACGAAATCCC CAACTCTCTC CGTGGGTCCT TGAGTTGCGT GAAATTTCTA  
601 GGATCACAAA TCAAAATGCA ATAAAATATG ATATGCATGA TGATCTAATG  
651 TATAACATTC CAATTGAAAA TTTGGGATGT TACATATAAC TCAAATTCTA  
701 TAATTATGAA CACAGAAATA TTAATGTAGA ACTCTATTTT GTTTTGAAAT  
751 TGTATTATTT TTAGAATTA GTCTAGAGCA TTTCGTGAAC TTGAATCAAA  
801 CCTTTAAATA AAACAAAGCA TAAAAATGAC AAATTCACAT ATGAAATAAC  
851 TTGTGTTACA TAGATTTATT ACAATAGCGT TGTATGTGTG TATGTGTGCG  
901 TGAGTGCCTA TGGTAATATC AATAAATATC TTGATAGATG TTTCTACAAT  
951 TCACGGGTCT AACTAGTAAT GCAATGCAAT GCATGCTAAA AGAATAGAAC  
1001 CTTAGTTTCA TTAACTAAC AATTTTCAAA TGTATGAGTT GCCAACAAGT  
1051 GGCATACTTG GCACTGTTTG TTTGTTTATT TTATGGAAAG TTCTTCTCTT

FIGURE 13a

1101 TTTACATGGT TTAGATTCCA GCATGTAGCC ACAAATATG ATTGTCAAAA  
 1151 GATAATACCT CATAATACAA TTCCACTAAA GTCACCTAGC CCAAGTGACC  
 1201 GACCTGATCC TGAAATAAAA TCAGAAGATT TGGTGTGATC ATCATGACAA  
 1251 CAAATTATTA GCGGGTAGAT CTTGTGGTAG TACTCATGAT GTAAAATTAT  
 1301 CAAGAGGGAG AGAATGTATG GAGATTTATG TGAAGTACAT CGTACACCAG  
 1351 ACATAGTTGA CACATCGATT TTTTAAGATA CATTTGGACG CGCCTTGTGG  
 1401 GAGTGTAAG TACTACCATG TATTAGAAGA GGTGAAATGA GAAATGCCAT  
 1451 AGCTAGCAAG TAGGCCTAGT TAAGGAAATT CTTCCTTAGA NTCCCCTTCT  
 1501 CCCGAAGAGT GAAGTGCTTC AACTAAAGGT TAGACCCACT TAAAAAATGT  
 1551 CACTTTGAAT CTTTGCTTCC CTTGTCGTAA TCCTGTGCAT TTGTAGGTCC  
 1601 CTCGGATCTG AGCCCTTTCT CCAAGCCCTT CATTGGATTC CCCTGGATGT  
 1651 CTTTTTGTTA CATTTTATTG AAGTGAGAGT GAATTATTAT ATGCCCATAG  
 1701 GAGGTGGGAT ATAAAGGCTG TTGGTATTCT GCACCATACA TGCTAGAGTA  
 1751 GGGAGGAGAG GCTGGTGCAT GATACATGGT GGACTAGCCC ATATATTTAC  
 1801 CCCTCCCCCA CCCACNTAAC AAGTTTTTTT NTATTAGGTC TTCATCCTCT  
 1851 GATTGTGTTT TCTGTTAGCC CATTCTTCAT CATGGACTTA TTAATCATGA  
 1901 TTAGTTTCTT GGATTTTGT TTAATTGACT TGAATTTGAC AATGTGCCTC  
 1951 ATATATGGCA TGTGGGACTG ATAGGAAGAT ATATTCTCAC AACATTA ACT  
 2001 TAAAAAGGAT TATTTTTTTG GTGCAGTCGT AAAGAAACT ACTTTCTTTT  
 2051 ATGCTAAAAG TTATTCAAAC ATAGATTTAT AAACAAAGGA TATCACCATG  
 2101 CATGACCATG CGCTCTCTCA TGTTTACTCT AGAAACCATA TATCTCTTTG  
 2151 TTGCAAAATA TTTAATCTAT CCTCCTTGTT TCTGGGAATG AGTCGGGGAA

FIGURE 13a (cont.)

2201 GGTAATCTTA GGGAAGGTTA AAGTGAGGCA AGTAAGAGCA ACTCTAGCAG  
 2251 AGTCGCGATA TGCCCAATCG CCATAATGCC AATATGGCAT TTTTGGCCCA  
 2301 AAATGGCACT TCAGAAGAGT CACCATATCC CTTCGGATAG CCATAATTTA  
 2351 GGGAGCTCGC TCCACAAACA AGCTTCGAGC CTCCAAATAT GGAGGCCATG  
 2401 GATTTCGTTGT TTGGCACTCA CTCCATATCC AACCGCAAGC GCATGCATGA  
 2451 GGGAAGTTTT AGCTTCTTCC TCCTTGCGCC AACGCCGGGA TTTTACACAG  
 2501 CGCATTACAG GTACATGAAC CAGCATGCAC AGATAATCAC CGACGAGTGG  
 2551 GGTGACAAGA AGGATAAGCA CCCTCCCATT AGTGGTGCGC CCACTCCCCT  
 2601 CAAATTCATG AGGCAGCCAT TTGGATGGTC ATCGCGTGGC ATAAGCTCCG  
 2651 ACTATAAAAT CTCAACGGCA TCACCAAAAC CATAGCTGCC GCCTCCCCCT  
 2701 TCCTCGGCAT CACCTCCCCA AGACATCTCC TCCCCTCTAT GCCACAATGT  
 2751 CATCATTATG GAGAGACACA ACNTACTGGN TAAACCGCAT ACCCAATCAT  
 2801 GGTTTACCGG CAGTGCGAAC CCCACCTTCC TCCCACGATG GTAGGATATT  
 2851 CTCCTCCTAG AATGGCGCGT GTGGCGCTTC CTCCTCCCGA GGCTGATATG  
 2901 TCGGCTCCCA TGATGGCGTG CATCATTGAT TTGGCGCTTC GGGTCCATCA  
 2951 TACATGTAA CGAGGTCATC CCCATTGATG TCGTTGGTCC CCTTGCCCCC  
 3001 CAGTCGGATC CTGAGGACCC GTTCGATGTC GCAATGCGAC TCTCCAAACT  
 3051 CAAAGCTCAC AATGAGGAGT ACGTCCTCTA GGAGTTCCGC CCCGCAACCA  
 3101 TCTATAAGGA GGAGCAACGA TAGCTCTCCC CTACGCCTTC CTCGACGATC  
 3151 TCTCTTAGGA GGACAACGGC TAGACGACGG CGGCGGCGGC GAAGGTACTG  
 3201 CAGGTAGTAG AACATAGCAA TGTCGAATGG CGACATTGCA TATTTTGAAA  
 3251 ATGTCGCTCA ACGACTTTTG AAGTCGCAA TAAAATGTAG TGTGACTACT

FIGURE 13a (cont.)

3301 TTTGGCCAGC AATATAAGTT TATCACATTT GATAATGATT TGAACCGGTG  
 3351 TGGTTCAACT AAATGTACCA TAAATTGAAC ATACAAATTT TTAGCAAATG  
 3401 AAAAAAGAAA CAAGTAAGAC CACAAATATG AAAGCCGCAT ATCGCGACTA  
 3451 TGTGTTTGAG CCGCAGCTGC CAAGTACATA TGAAGCGTAC TCCATATGAC  
 3501 ATACGACAAC CATACATATG AAGACTCTAC TAGAGTTCTC TAAGGCCGCT  
 3551 TTTAGCGCCT TTCGTGCAGT GGTGCCATA GGGAGTGAGG GTAGTTGGAC  
 3601 TGTTTCGTTT CCCTTTTTTC ATTTCTTTGA AATCTATTTT ATTTTTTTTC  
 3651 TCTTTTGTAG GTTCCCAA TTTATATACC ATTTTCTGT TTCTCGCTAT  
 3701 TTTTGTGTGT TATATTCTAG TTTCATATTT TTCTATTATT AATTTGTGTC  
 3751 TCTTATGAGA AGTCCAGACT TGCATATGGA GGTGCACACA CAAACATATA  
 3801 AAGTATAAAT ACTAACTTGA GAAGTATGTT TGC GTGGTCA AAAAAACATC  
 3851 ATCAAAACCT GCCAATATGA GATATAGTTT TGAATATATC AATATGAGCA  
 3901 ACGCAACCAT TTAAAATGTG AACAATTGTT TTTT TAGAAA AAATATAAGA  
 3951 AATAACTCCA ACCCAGCCAA ACCACATGCT ATACACTTGC TCCATATGAA  
 4001 ACCATGTTTG CTATTGGGCA GTTGCCTGAA ACCGAAAGTA ATGTTAGCCG  
 4051 TTTTCTATT CAAAGAAGAA GGAGAGTCGA GGTGACGCGA TGCTTAGACG  
 4101 NTGAGATGGG GATGACCACA ACGTCCCTAC AGAGACCTCA CCGGAGATGG  
 4151 GGACATTGCA GTTGACACGA GAGCGGTGAG GGGCTGCGAT GCGTGTGCGG  
 4201 CAACATGTGG CGAGGCGGAC GTCGGGCTGG CAGGTAGGGG GGAGGGGGAA  
 4251 GGACCGGGGG AGGAAGAAGA GGAGTAGCCT GCAAAACATG GTACACCAGT  
 4301 TTTCTGCCCT ACGAAAACCT CATTTTCATTC CCCCACCCTG ACAAGCAACA  
 4351 ACCAACCATC GCAGTCCCAC ATGTCCCTCT GGTCTTTGCA AAAAGTAATT

FIGURE 13a (cont.)

19/66

4401 GTTCTTGCTG GACAGCGCAA AGAGTAAACT TTTGTTAGTT TTCATTTCTA  
4451 GAAAAAGCAA TCCTTTTATA GTTCTTTTGT GAAAGTAATG CTTTATAGT  
4501 GATTGGGATG TTCTTTTAGA GCAAATATCT TCTTTTTTTT TTAGGGAAAA  
4551 GAGCAAATAT CTTCCACTTT TCACAAAACT GACGAAGGCT GAAAGTGGCG  
4601 AGACANGTGA GGGCCCATAG CTTTCGTCCG GCCCAGCGGC GCACGACCGT  
4651 CCACGTGCAC CCCGGCCCTC CCGGGCCCGC AGATCCGNTT CTCCTCGCC  
4701 CCCGTTTCCC CCTCCCTCCC TCTCGTTGCT TCCACTCCAC TGTTCTCCTC  
4751 TTCCTGTCCA AAGCGGCCAC GGACCGGAAA AAAATCACGC CTTTCCGTTG  
4801 GGTCTCCGGC GCCACACTCC TCCTCCGGCC GATATAAAGC GCGCGGGGCC  
4851 ACGGGCCCGG CGCAAAATGG GATTCCCGTC CGCCGCCATG GAGGAAGATG

FIGURE 13a (cont.)

20/66

1 ACGGGCCCGG CGCAAAATGG GATTCCCGTC CGCCGCCATG GAGGAAGATG  
51 TTCTGCTTCA CCGCCCCTTC CTGNTCGCCA TNTCTCCCGC CGCGCCCNTC  
101 CCGTCCCGNT GCTGACCGGC CCGGACCGGG GATCTCGGTG AGTCAGTCGG  
151 GATCTTCATT TCTTTTCTTT TCTTTCGTTT CCGGCNTCCG TTCTGCCGGG  
201 GTTTCCTGA TGCATGCCG CGCGCGCGCA GGGCGGCGGC AATGTGCGGC  
251 TGAGCGCGGT GCCCGCGCCC TCTTCGCTCC GCTGGTGTGG CCGCGGAAGG  
301 TGAGCCCTCT CCCCTGTCTA CCCAGATTTG CGACCGTGAT CCCCTGTTGT  
351 CGCCGGGCAA ACGGAATCTG ATCCACGGTG GTTATTGGAA ATAGTATATA  
401 CTACTAATAA ACTTGAGGCT GGGATTCGTC CACTGAGGAA CAAGTGGATG  
451 CGATTTTCGAT TGGATTTCTC TGCTTTATGC GATCCGTACG CAGAATATCC  
501 CTCCTGCAGT GTCTCAACCG TATTACTGGA TGTACAACCC AAATGTGTAT  
551 AATCTGTGCT GAATGTATCA ACCAATAATT GCTGCATTGT GAAAACATAA  
601 TCCTGTGTTG TGTCTCTACT ACTTGTTTCTC TCCTGATCTG CCGCTTATCC  
651 TAACTTTTGT TCATTTATGG AAGGCCAAGA GCAAGTTCTC TGTTCCTG  
701 TCTGCGCCAA GAGACTACAC CATGGCAACA GCTGAAGATG GTGTTGGCGA  
751 CCTTCCGATA TACGATCTGG ATCCGAAGTT TGCCGGCTTC AAGGAACACT  
801 TCAGTTATAG GATGAAAAAG TACCTTGACC AGAAACATTC GATTGAGAAG  
851 CACGAGGGAG GCCTTGAAGA GTTCTCTAAA GGTTAGCTTT TGTTTCATGT  
901 GTTTGAAACA ATAGTTACAT CTTGTGGCGT CCGCAGCACA AAAGACATAA  
951 TGCGACTCTG TTTTGTAGGC TATTTGAAGT TTGGGATCAA CACAGAAAAT  
1001 GACGCAACTG TGTACCGGGA ATGGGCCCCT GCAGCAATGT AAGTTCTAGT  
1051 GTTGTACACG AACTAATTGC AATGGTCGTT GGTAACTTA TGAAGTGCTG

FIGURE 13b

1101 ATGAAACTGT CTTAAGAGTT TATGGCTTGT CTTTTCTGAT TCTAGCTAGT  
1151 AAAGAGTAGA TAAATATGAA ATATGTTTTTC CCTTTTCTAG TTATGGTCAT  
1201 GGTGGCTGG TATTCATTTTC TTTTATGGCA ATACTTGCTT CTAACATATCT  
1251 TTAGTAGATT CATGTATTTA CTTGTGAGTC ATTACTTTAT GGGTGTAGGG  
1301 ATGCACAACT TATTGGTGAC TTCAACAACT GGAATGGCTC TGGGCACAGG  
1351 ATGACAAAGG ATAATTATGG TGTTTGGTCA ATCAGGATTT CCCATGTCAA  
1401 TGGGAAACCT GCCATCCCCC ATAATTCCAA GGTAAATTT CGATTTACC  
1451 GTGGAGATGG ACTATGGGTC GATCGGGTTC CTGCATGGAT TCGTTATGCA  
1551 TCCACCTTCT GGTGAAAGGT CTACTTTTAG TGGCTCGAGA GCAAGAAATC  
1601 TAAGTAAAC CCACACAATT AACTTACATT AATGTGGAGA CATGATACTT  
1651 TTATTGCTCG TTTTGCAGGT ATGTGTTTAA GCATCCTCGG CCTCGAAAGC  
1701 CTGACGCTCC ACGTATTTAC GAGGCTCATG TGGGGATGAG TGGTGAAAAG  
1751 CCTGAAGTAA GCACATACAG AGAATTTGCA GACAATGTGT TACCGCGCAT  
1801 AAAGGCAAAC AACTACAACA CAGTTCAGCT GATGGCAATC ATGGAACATT  
1851 CCATATTATG CTTCTTTTGG GTACATGTTG ACGAATTTCT TCGCAGTTAG  
1901 CAGCAGATCA GGAACCCAGA AGACCTCCAA TATCTTGTTG ACAAGGCACA  
1951 TAGTTTAGGT TGC GTGTTCT GATGGATGTT GTCCATAGCC ATGCGAGCAG  
2001 TAATAAGACA GATGGTCTTA ATGGCTATGA TGTTGGGCAA AACACACAGG  
2051 AGTCCTATTT CCACACAGGA GAAAGGGGCT ATCATAAACT GTGGGATAGC  
2101 CGCCTGTTCA ACTATGCCAA TTGGGANGTC TTACGATTTT TTCTTTCTAA  
2151 TCTGAGATAT TGGATGGACG AATTCATGTT TGATGGCTTC CGATTTGATG

FIGURE 13b (cont.)

22/66

2201 GGGTAACATC CATGCTATAT AATCACCATG GTATCAATAT GTCATTGCT  
2251 GGAAGTTACA AGGAATATTT TGGTTTGGAT ACTGATGTAG ATGCAGTTGT  
2301 TTACCTGATG CTTGCGAACC ATTTAATGCA CAAACTCTTG CCAGAAGCAA  
2351 CTGTTGTTGC AGAAGATGTT TCAGGCATGC CAGTGCTTTG TCGGTCAGTT  
2401 GATGAAGGTG GAGTAGGGTT TGACTATCGC CTGGCTATGG CTATTCCTGA  
2451 TAGATGGATC GACTACTTGA AGAACAAAGA TGACCTTGAA TGGTCAATGA  
2501 GTGGAATAGC ACATACTCTG ACCAACAGGA GATATACGGA AAAGTGCATT  
2551 GCATATGCTG AGAGCCATGA TCAGGTATGT TTTCCCTCCT TTGTCGCTGT  
2601 GCGTGAGTAT GTGTTCTTTT TTTATGGGGC ACTGGTCTAA GAACATACAG  
2651 TTCAAAGGTG AGACACTTTC TTTGCCTGGT AGACAAATTT GAGAAATAAA  
2701 CATTTGCTT GATGACTTTT AGTTGCTTCA CAAGTTCGAA TTAAGTTAGT  
2751 TATATTCTGA TAACTAGTGA TAGTACCCAC TAACCAGCTA TTACGGACCA  
2801 TGTAAGAATG TCCGAAGACT GCAGTTATAT ATCGTTGACT TTGTGTTTAT  
2851 CTATTGAAAC AACTTAGTAG TTAACCTTCA CGCAAATTTT CAGTCTATTG  
2901 TTGGCGACAA GACTATGGCA TTTCTCTTGA TGGACAAGGA AATGTATACT  
2951 GGCATGTCAG ACTTGCAGCC TGCTTCGCCT ACAATTGATC GTGGAATTGC  
3001 ACTTCAAAAG GTTCGATTCTG TTTTAAGTAT TCCTGAATTT GATGTTCTAG  
3051 TTCCAGACGA GTATTGTAAT GTTCGTTGTT ACTCAGAGTT CTGCTTAGTC  
3101 CTTGAAGATA ATGTATTCCA GTCCCTTTTG GTACATTTGG CTTATTTTGT  
3151 TACAAATATT TCAGATGATT CACTTCATCA CCATGGCCCT TGGAGGTGAT  
3201 GGCTACTTGA ATTTTATGGG TAATGAGGTA ATATCTGGTT ATCTGTCAAA  
3251 ACTTATTTCT GATCAATATG TTTCGGGATT CCCTCGAAAA AAATCCTTTG  
3301 GGCAGGGCGA AAAGTTTAAA CATCTGTTTT CTATGATAGC CAAGTACTCC  
3351 CCAGCTATTT CCATGTTATC ACGTATCATT TAGCTGTGCC GGTAGTTAAT

FIGURE 13b (cont.)



3401 CTTTATTCTA ATTCATTGTT GTTTTTTAGC GTGGCAGTCT ATTGTTGGAT  
 3451 CCTCTTATTC CAATTACATA TATGCCGACA TCACACACTT ATGAATATTC  
 3501 CCTGTTTAAA AGATTTTTTAT TTTATACCAA TGTTTCTCCG TAAATGATGC  
 3551 AAACATGATA GAGATGTTAG CATGTCTTTC TTAACCTACT CATGTTTTAC  
 3601 ATATCACGAC AAGCTTCTTG CAGAAAATCA GCAGTATATG GCAAATTGCT  
 3651 GCAACCTGAC AACGTTTATA TCTGTTTTCT AACTCATACT GACGGTGCAA  
 3701 TTTCTTTTTA GTTTGGCCAC CCAGAATGGA TTGACTTTC AAGAAGAAGG  
 3751 CAACAACCTGG AGTTATGATA AATGCAGACG CCAGTGGAGC CTCGCAGACA  
 3801 TTGATCACCT ACGATACAAG GTTATGCCTA TGTATATTTT TACAGTTTCT  
 3851 GGTCTGGTAG CTCTCTTGGG ATCTTGACCT CACTTAGTTC CTTCATCTCT  
 3901 GACTGTAGCT TATTTACACT GTGTTCCAAC TTCTGTCTTG TGGATAAATT  
 3951 CTCCCTTCTA ACGTTTCATA TTAAGCCTTT CAAACTAAAC TAAATTGCTG  
 4001 ATCTACTACT AGTTGCTCAG TACGATGACC AAATCTTGCC TGTGGTAACC  
 4051 TAGTAATTTT CTTGATTCTT ACACATTAGT GATATGCAGT GCATACATTA  
 4101 TCCATATAAA TTGACATTGC AATTTCCCAA ATATTATTIG AAGGCTGTGT  
 4151 TCTTTTGTGA ACAGGAAGTT ATTTTCTCTG CATCTGATAA ATAATAATAG  
 4201 CCTTTCACGA TTTTCTCAT ATTTTATCCA ACTTTTCTGC ATTCAAGCAT  
 4251 TTTTTGTTTC TCGCCTAACA TATATAATTT GAACAGTACA TGAACGCATT  
 4301 TGATCAAGCA ATGAATGCGC TCGACGACAA ATTTTCCTTC CTATCATCAT  
 4351 CAAAGCAGAT TGTCAGCGAC ATGAATGAGG AAAAGAAGTA GTTAACTATA  
 4401 CAATGTTTAG TCAGGGCAGC TGTTGCATCA TTTGATTAC TCCTACTCTT  
 4451 AAGAATAGCA ACTCTGACTT GTGCGTTTTA TGTTACCAA TAAGTTGAAA  
 4501 CCGTATCTGT TTGATATGAA CCATTGTTGT CTCAAATGG GCTATGGACT  
 4551 CAATCCAAC TCCTTTCAG ATTATTGTAT TTGAACGTGG ANATCTGGTC

FIGURE 13b (cont.)

4601 TTCGTCTTCA ATTTTCATCC CAGTAAAACT TATGATGGGT AACTGATCTC  
 4651 TTGCAAGCTT TGCCTTTCAA TATTTCTTCT GCTTAATGAC TAATGTGCTT  
 4701 AATCTCGTTT CCACTTTTAA AACACGCAGT TACAAAGTCG GATGTGACTT  
 4751 GCCTGGGAAG TACAAGGTAG CTCTGGACTC TGATGCTCTG ATGTTTGGTG  
 4801 GACATGGAAG AGTAAGCAAT GTTAATGATG TTCAAGATCT GTTTTGCAAC  
 4851 ACTATGTTCT TCTATAGAAG GGGCCATCAA GGCTGCATCA GATAATCTTA  
 4901 TTTGCAGTGT TGATCTGTGC TGCATCGCAG GTGGCCCATG ACAACGATCA  
 4951 CTTTACGTCA CCTGAAGGAG TACCAGGAGT ACCTGAAACA AACTTCAACA  
 5001 ACCGCCCTAA CTCATTCAA ATCCTGTCTC CATCCCGCAC TTGTGTGGTA  
 5051 ATGCTAATTA CTAGGAGGAT TTAGTAACAA TAAATAAATA ACAGCAAAAG  
 5101 ATATCTGCAG TACGATCTCA CAAAATGCTC TCTTGCCAGG CTTACTATCG  
 5151 CGTCGAGGAG AAAGCGGAAA AGCCCAAGGA TGAAGGAGCT GCTTTCTTGG  
 5201 GGGAAACTGC TCTCGGGTAC ATCGATGTTG AAGCCACTGG CGTCAAAGAC  
 5251 GCAGCAGATG GTGAGGCGAC TTCTGGTTCC GAAAAGGCGT CTACAGGAGG  
 5301 TGACTCCAGC AAGAAGGGAA TTAACTTTGT CTTTCTGTCA CCCGACAAAG  
 5351 ACAACAAATA AGCACCATAT CAACGCTTGA TCAGGACCGT GTGCCGACGT  
 5401 CTTTGTAATA CTCCTGCTAT TGCTAGTAGT AGCAATACTG TCAAACGTGT  
 5451 CAGACTTGAA ATTCTGGCTT GGACTTTGCT GAGGTTACCT ACTATATAGA  
 5501 AAGATAAATA AGCGGTGATG GTGCGGGTCG AGTCCAGCTA TATGTGCCAA  
 5551 ATATGCGCCA TCCCGAGTCC TCTGTCATAA AGAAAGTTTC GGGCTTCCAT  
 5601 CCCAGAATAA AAACAGTTGT CTGTTTGCAA TTTCTTTTTG TCTTGCATAG  
 5651 TTACATGATA ATTGATGCAT ATTGCTATAA GCCTGGATTG CATCTTCTTT  
 5701 TGCTAATAAC TGCAGGGCCA AGAAAGCCTA GATTGTATCT TTTTTTGCTA  
 5751 ATAAC TGCAG TGCTGGGGAA GCTTCAGTCC TTGTTTCCGT TCTCGAGACA

FIGURE 13b (cont.)

25/66

5801 AGGCGTCATG TTTGGCGCAC AAAGGTAAGC CATCATCTTA TCAAGTCCCA  
5851 AAATTCTCTG GTTGAAAGAA ACCATCACTA ACTTGTTCCA GGTGTTGGTT  
5901 CCTCCACAAC CAAAAGGCGA CCATCGTCGT CATCATCGCT CACAGCACTG  
5951 ACCATCGAAG CCACGGTGGG CATGANAANT GCGCATCGCC CAAGACTTGG  
6001 GACCGTTTCA AAANTATCAC AAAGTGCCAT GGNCATCTTC TGCCAAAGGC  
6051 TGCACTGCAC CTTTGGCATG AACAGAAGCA ANNCAGGGGC TTGGAAGTGA  
6101 ACNGCCGAAA ATAAAGTCAA NACCGGCTGG GCCGGATTGA AAGGGGAAAC  
6151 GNCCAAAATC CACTTNAATT TGAATGGAAG GANGGAATGG TTCTTGCTGG  
6201 TTTTCAACTC TGCANGGCTT CCNCTCTGAA TTTACACGCG ANGNCCATT

FIGURE 13b (cont.)

26/66

1 GCGACTTCTG GTTCCAAAAA GCGTCTACA GGGAGGTGAC TCCAGCAAGA  
51 AGGGAATTAA CTTTGTCTTC GGGTCACCTG ACAAAGATAA CAAATAAGCA  
101 CCATATCAAC GCTTGATCAG AACCGTGTAC CGACGTCCTT GTAATATTCC  
151 TGCTATTGCT AGTAGTAGCA ATACTGTCAA ACTGTGCAGA CTTGAGATTC  
201 TGGCTTGGAC TTTGCTGAGG TTACCTACTA TATAGAAAGA TAAATAAGAG  
251 GTGATGGTGC GGGTCGAGTC CGGCTATATG TGCCAAATAT GCGCCATCCC  
301 GAGTCCTCTG TCATAAAGGA

FIGURE 14

27/66

Genomic clones from *T. tauschii*  
for SBE-II

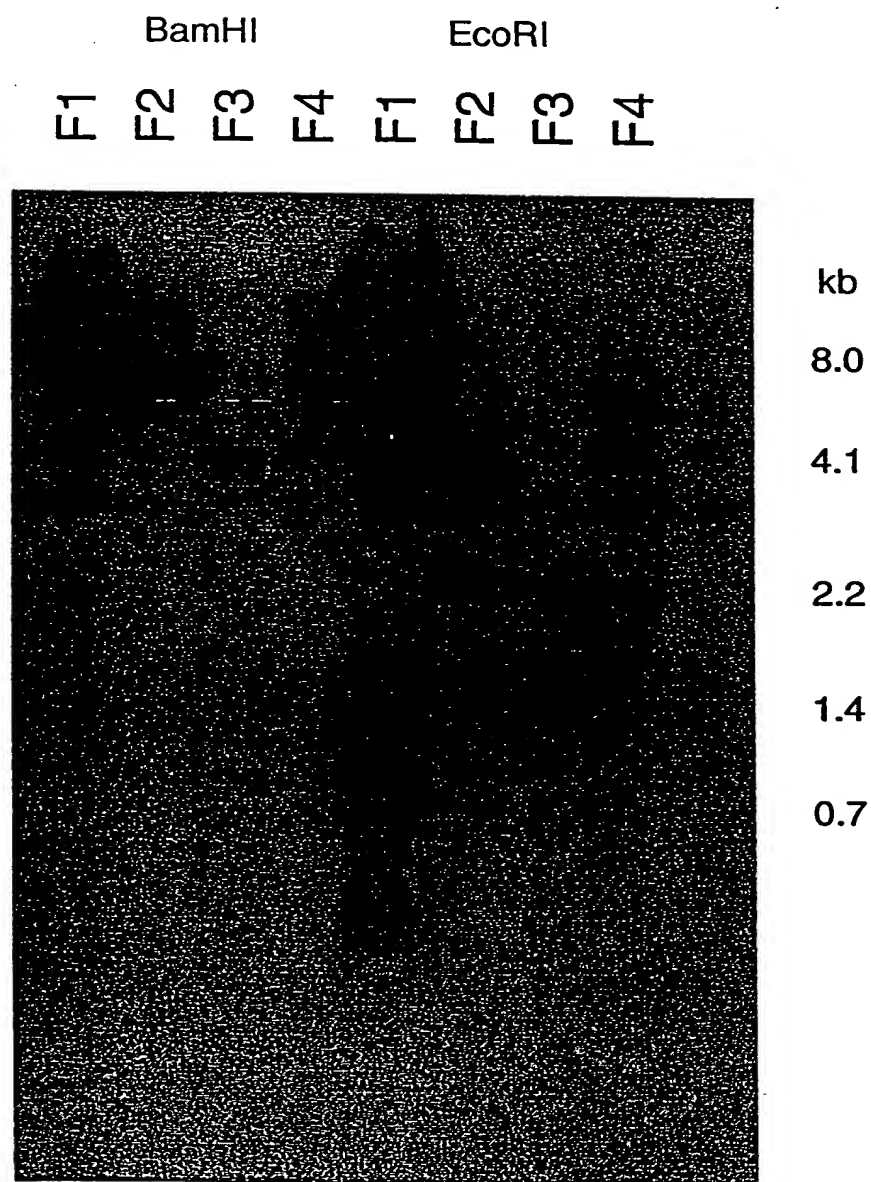


FIGURE 15

*N-terminal sequences of cereal starch branching enzymes*

Protein	1	2	3	4	5	6	7	8	9	1	1	1	1	1	1	1	1	2	2	2
	A									0	1	2	3	4	5	6	7	8	9	0
RICEBEI <sup>B</sup>	A	T	A	R	K	N	K	T	M	V	T	V	V	E	E	V				
WBE-I <sup>AD</sup>	V	S	A	P	R	D	Y	T	M	A	T	A	E	D	G	V				
MAIZE	A	T	V	Q	E	D	K	T	M	A	T	A	K	G	D	V				
BEI <sup>C</sup>																				
RICEBEII <sup>D</sup>	A	A	G	A	S	G	E	-	V	M	I	P	E	G	E	S	D	G	M	P
WBE-II																				
MAIZE	A	A	A	A	R	K	A	V	M	V	P	E	G	E	S	D	D	L	A	S
BEII <sup>E</sup>																				

<sup>A</sup> N-terminal amino acid of the mature polypeptide. <sup>B</sup> Kawasaki *et al.* (1993), <sup>C</sup> Babu *et al.* (1991),

<sup>D</sup> Mizuno *et al.* (1993), <sup>E</sup> Fisher *et al.* (1993)

Residues in the wheat sequences showing identity with the respective maize or rice branching enzyme isoforms are highlighted in bold text.

FIGURE 16a

29/66

1 TTCCCTTTTTTTTCTTTGGGNGGGGGAATGCCCTGTGGATGNTGTTCCCAATGAATTT 60  
 AAGGGAAAAAAGAAACCCNCCCCCTACCGGACAACTACNACAAGGGGTTACTTAA

a F P F F F F G ? G M A C W M ? F P N E F -  
 b S L F F S L G G G W P V G ? C S P M N F -  
 c P F F F L W ? G D G L L D ? V P Q \* I S -

61 CCATGGAGTGAGAGAGATAGTTGGATNAGGGATCGCGNTTCCNGGAACGTATTTTTTTC  
 GGTACCTCACTCTCTCTATCAACCTANTCCCTAGCGCNAAGGNCCTTGACATAAAAAAG 120

a P W S E R D S W ? R D R ? S ? N C I F F -  
 b H G V R E I V G ? G I A ? P G T V F F S -  
 c M E \* E R \* L D ? G S R F ? E L Y F F P -

121 CCCNGCGGGGAAATGGCGTTAGTGTCTNACCCAGGCCCTGGTGTACCACGGCTTTGATC  
 GGGNCGCCCCCTTTACCGCAATCACAGNTGGGTCCGGGACCACAATGGTGCCGAAACTAG 180

a P ? G G N G V S V ? P G P G V T T A L I -  
 b P A G E M A L V S T Q A L V L P R L \* S -  
 c ? R G K W R \* C ? P R P W C Y H G F D H -

181 ATTCTTCGTTTCATTCTGATATATATTTTCTCATCTTTTCTTCTCTGTTCTGCTGTAA  
 TAAGAAGCAAAGTAAGACTATATATAAAGAGTAAGAAAAAGAAGGACAAGAAGACATT 240

a I L R F I L I Y I F S F F F F L F L L \* -  
 b F F V S F \* Y I F S H S F S S C S C C N -  
 c S S F H S D I Y F L I L F L P V L A V T -

241 CTGCAAGTTGTGGCGTTTCTTCACTATGTAGTCATCCTTGCATTTTGCAGGCGCGTCC  
 GACGTTCAACACCGCAAAAAGTGATAACATCAGTAGGAACGTAAAAAGTCCGCGCAGG 300

a L Q V V A F F H Y C S H P C I L Q A P S -  
 b C K L W R F F T I V V I L A F C R R R P -  
 c A S C G V F S L L \* S S L H F A G A V L -

301 TGAGCGCGCGGCGCTCTCCAGGGAAGGTCTGGTGCTGACGGGAGAGNGAAGACTTGG  
 ACTCGGCGCGCGGAGAGGTCCCTTCCAGGACCAAGGACTGCCGCTCTCCTGCTGAAC 360

a \* A A R P L Q G R S W C L T A R ? T T W -  
 b E P R G L S R E G P G A \* R R E ? R L G -  
 c S R A A S P G K V L V P D G E ? D D L A -

361 CAAGTCCGGCGCAACCTGAAGAATTACAGGTACACACTCGTGCCGGTAAATCTTCATA  
 GTTCAGGCCGCGTTGGACTTCTTAATGTCCATGTGTGTGAGCAAGGCCATTTAGAAGTAT 420

a Q V R R N L K N Y R Y T H S C R \* I F I -  
 b K S G A T \* R I T G T H T R A G K S S Y -  
 c S P A Q P E E L Q V H T L V P V N L H T -

421 CAATCGTTATTCACTTACCAAATGCCGGATGAAACCAACCAGGATGCGTCAGGTTTGA  
 GTTAGCAATAAGTGAATGGTTTACGGCCTACTTTGGTTGGTGCTACGCAGTCCAAAGCT 480

a Q S I F T Y Q M P D E T N H G C V R F R -  
 b N R Y S L T K C R M K P T T D A S G F E -

FIGURE 16b

30/66

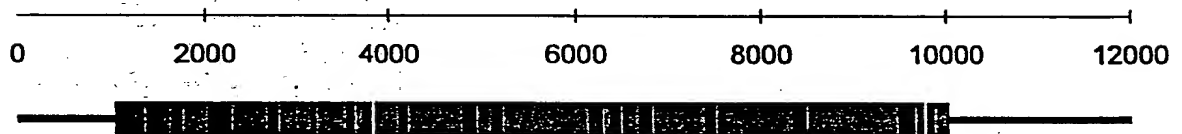
1 MATFAVSGAT LGVARPPAAA QPEELQIPED IEEQTAEVNM TGGTAEKLES  
51 SEPTQGIVET ITDGVTKGVK ELVVGEKPRV VPKPGDGQKI YEIDPTLKDF  
101 RSHLDYRYSE YRRIRAAIDQ HEGGLEAFSR GYEKLGFTSR AEGITYREWA  
151 PGAHSAALVG DFNNWNPAD TMTRDDYGVW EIFLPNNADG SPAIPHGSRV  
201 KIRMDTPSGV KDSISAWIKF SVQAPGEIPF NGIYYDPPEE EKYVFQHPQP  
251 KRPESLRIYE SHIGMSSPEP KINSYANFRD EVLPRIKRLG YNAVQIMAIQ  
301 EHSYYASFGY HVTNFFAPSS RFGTPEDLKS LIDRAHELGL LVLMDIVHSH  
351 SSNNTLDGLN GFDGTDTHYF HGGPRGHHWM WDSRLFNYS WEVLRFLLSN  
401 ARWWLEEYKF DGFRFDGVTS MMYTHHGLQM TFTGNNGEYF GFATDVEDAVV  
451 YLMLVNDLIH GLHPDAVSIG EDVSGMPTFC IPVDPGGVGF DYRLHMAVAD  
501 KWIELLKQSD ESWKMGDIVH TLTNRRWLEK CVTYAESHDQ ALVGDKTIAF  
551 WLMDKDMYDF MALDRPSTPR IDRGIALHKM IRLVTMGLGG EGYLNFMGNE  
601 FGHPEWIDFP RGPQTLPTGK VLPGNNSYD KCRRRFDLGD ADFLRYHGMQ  
651 EFDQAMQHLE EKYGFMTSEH QYVSRKHEED KVIIFERGD L VFVFNHWSN  
701 SFFDYRVGCS RPGKYKVALD SDDALFGGFS RLDHDVDYFT TEHPHDNRPR  
751 SFSVYTPSRT AVVYALTE\*

FIGURE 16C



*Branching Enzyme-II Genes*

## Intron/Exon structure of wheat BE-II



## Schematic Diagram of a cDNA for BE-II

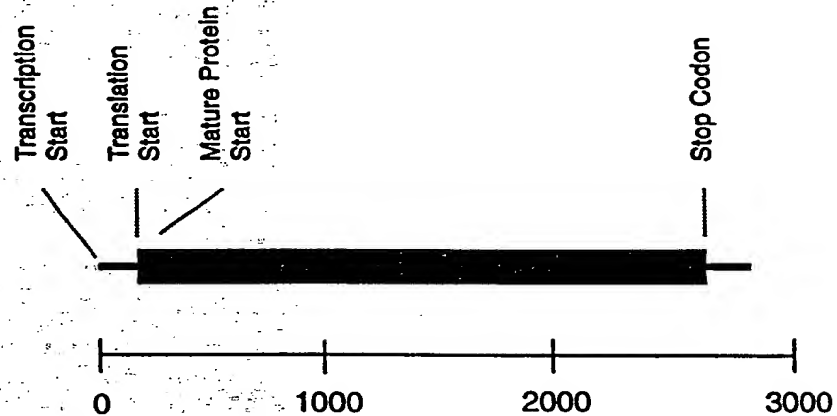
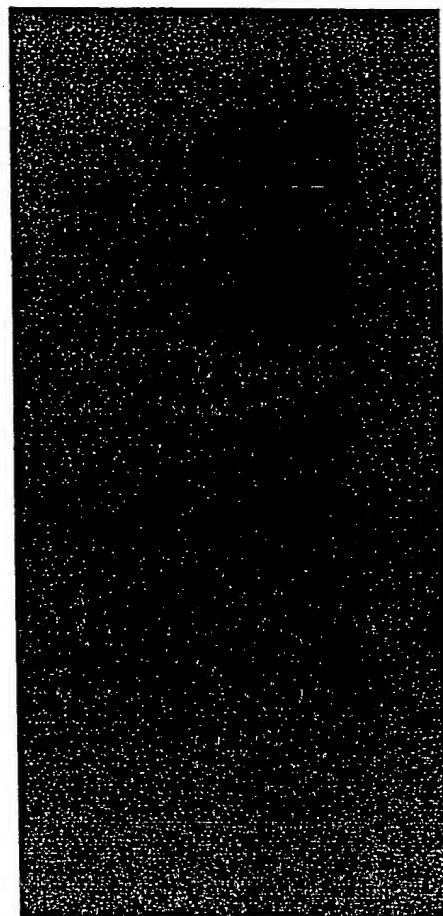


FIGURE 17

32/66

Wheat DNA Probed  
with 5' end of SBE-II

N2AT2B  
N2BT2A  
N2DT2A



8 kb

2.2 kb

FIGURE 18

```

1  AGAAACACCT CCATTTTAGA TTTTTTTTTT GTTCTTTTCG GACGGTGGGT
51  CGTGGAGAGA TTAGCGTCTA GTTTTCTTAA AAGAACAGGC CATTTAGGCC
101 CTGCTTTACA AAAGGCTCAA CCAGTCCAAA ACGTCTGCTA GGATCACCAG
151 CTGCAAAGTT AAGCGCGAGA CCACCAAAAC AGGCGCATTC GAACTGGACA
201 GACGCTCACG CAGGAGCCCA GCACCACAGG CTTGAGCCTG ACAGCGGACG
251 TGAGTGCGTG ACACATGGGG TCATCTATGG GCGTCGGAGC AAGGAAGAGA
301 GACGCACATG AACACCATGA TGATGCTATC AGGCCTGATG GAAGGAGCAA
351 CCATGCACCT TTTCCCCTCT GGAAATTCAT AGCTCACACT TTTTTTTAAT
401 GGAAGCAAGA GTTGGCAAAC ACATGCATTT TCAAACAAGG GAAAATTAAT
451 TCTCAAACCA CCATGACATG CAATTCTCAA ACCATGCACC GACGAGTCCA
501 TGCGAGGTGG AAACGAAGAA CTGAAAATCA ACATCCCAGT TGTGAGTGC
551 AGAAGAGGAT GACACTGAAA GTATGCGTAT TACGATTTCA TTTACATACA
601 TGTACAAATA CATAATGTAC CCTACAATTT GTTTTTTGGG GCAGAGTGGT
651 GTGGTCTTTT TTTTTTACAC GAAAATGCCA TAGCTGGCCC GCATGCGTGC
701 AGATCGGATG ATCGGTCGGA GACGACGGAC AATCAGACAC TCACCAACTG
751 CTTTTGTCTG GGANACAATA AATGTTTTTT GTAAACAAAA TAAATACTTA
801 TAAACGAAGG GTACTAGAGG CCGCTAACGG CATGGCCAGG TAAACGCGCT
851 CCCAGCCGTT GGTTTGCNAT CTCGTCCTCC CGCACGCAGC GTCGCCTCCA
901 CCGTCCGTCC GTCGCTGCCA CCTCTGCTGT GCGCGCGCAC AAGGGAGGAA
951 AACACGCCG CACACACACT CACACACGGN ACACTCCCCG TGGGTCCCCT
1001 TTCCGGCTTG GCNTCTATCT CCTCTCCCCC GCCCATCCCC ATGCACTGCA
1051 CCGTACCCGC CAGCTTCCAC CCCC GCCGCA CACNTTGCTC CCCCTTCTCA
1101 TCGCTTCTCA ATTAATATCT CCATCACTCG GGTTCGCGC TGCATTTCTG
1151 CCGGCGGGTT GAGTGAGATC TGGGCGACTG GCTGACTCAA TCACTACGCG
1201 GGGATG

```

FIGURE 19a

1 CCGGCGGGTT GAGTGAGATC TGGGCGACTG GCTGACTCAA TCACTACGCG  
 51 GGGATGGCGA CGTTCGCGGT GTCCGGCGCG ACTnTCGGTG TGGCGCGGGC  
 101 CGGCGTCGGA GTGGCGCGGG CCGGCTCGGA GCGGAGGGGC GGGGCGGACT  
 151 TGGCGTCGCT GCTCCTCAGG AAGAAGGACT CCTCTCGTAC GCTTCGCTCT  
 201 CTCGAATCTC CCCCGTCTGG CTTTGGCTCC CCTTCTCTCT CCTCTGCGCG  
 251 CGCATGGCCT GTTCGATGCT GTTCCCCAAT TGATCTCCAT GAGTGAGAGA  
 301 GATAGCTGGA TTAGGCGATC GCGCTTCCTG AACCTGTATT TTTTCCCCCG  
 351 CGGGGAAATG CGTTAGTGTC ACCCAGGCCC TGGTGTTACC ACGGCTTTGA  
 401 TCATTCCCTCG TTTCATTCTG ATATATATTT TCTCATTCTT TTTCTTCCTG  
 451 TTCTTGCTGT AACTGCAAGT TGTGGCGTTT TTTCACTATT GTAGTCATCC  
 501 TTGCATTTTG CAGGCGCCGT CCTGAGCCGC GCGGCCTCTC CAGGGAAGGT  
 551 CCTGGTGCCT GACGGCGAGA GnGACGACTT GGCAAGTCCG GCGCAACCTG  
 601 AAGAATTACA GGTACACACA CTCGTGCCGG TAAATCTTCA TACAATCGTT  
 651 ATTCACTTAC CAAATGCCGG ATGAAACCAA CCACGGATGC GTCAGGTTTC  
 701 GAGCTTCTTC TATCAGCATT GTGCAGTACT GCACTGCCTT GTTCATTTTG  
 751 TTAGCCTTGG CCCCGTGCTG GCTCTTGGGC CACTGAAAAA ATCAGATGGA  
 801 TGTGCATTCT AGCAAGAACT TCACAACATA ATGCACCGTT TGGGGTTTCG  
 851 TCAGTCTGCT CTACAATTGC TATTTTTTCGT GCTGTAGATA CCTGAAGATA  
 901 TCGAGGAGCA AACGGCGGAA GTGAACATGA CAGGGGGGAC TGCAGAGAAA  
 951 CTTCAATCTT CAGAACCGAC TCAGGGCATT GTGGAAACAA TCACTGATGG

FIGURE 19b

1001 TGTAACCAAA GGAGTTAAGG AACTAGTCGT GGGGGAGAAA CCGCGAGTTG  
1051 TCCCAAAACC AGGAGATGGG CAGAAAATAT ACGAGATTGA CCAACACTG  
1101 AAAGATTTTC GGAGCCATCT TGACTACCGG TAATGCCTAC CCGCTGCTTT  
1151 CGCTCATTTT GAATTAAGGT CCTTTCATCA TGCAAATTTG GGGAACATCA  
1201 AAGAGACAAA GACTAGGGAC CACCATTTC A TACAGATCCC TTCGTGGTCT  
1251 GAGAATATGC TGGGAAGTAA ATGTATAATT GATGGCTACA ATTTGCTCAA  
1301 AATTGCAATA CGAATAACTG TCTCCGATCA TTACAATTAA AGAGTGGCAA  
1351 ACTGATGAAA ATGTGGTGGA TGGGTTATAG ATTTTACTTT GCTAATTCCT  
1401 CTACCAAATT CCTAGGGGGG AAATCTACCA GTTGGGAAAC TTAGTTTCTT  
1451 ATCTTTGTGG CCTTTTTGTT TTGGGGAAAA CACATTGCTA AATTCGAATG  
1501 ATTTTGGGTA TACCTCGGTG GATTCAACAG ATACAGCGAA TACAAGAGTG  
1551 CTGCTATTGA CCAACATGAA GGTGGATTGG AAGCATTTTC TCGTGGTTAT  
1601 GAAAAGCTTG GATTTACCCG CAGGTAAATT TAAAGCTTTA TTATTATGAA  
1651 ACGCCTCCAC TAGTCTAATT GCATATCTTA TAAGAAAATT TATAATTCCT  
1701 GTTTTCCCCT CTCTTTTTTC CAGTGCTGAA GGTATCGTCT AATTGCATAT  
1751 CTTATAAGAA AATTTATATT CCTGTTTTCC CCTATTTTCC AGTGCTGAAG  
1801 GTATCACTTA CCGAGAATGG GCTCCCTGGA GCGCATGTTA TGTTCTTTTA  
1851 AGTTCCTTAA CGAGACACCT TCCAATTTAT TGTTAATGGT CACTATTCAC  
1901 CAACTAGCTT ACTGGACTTA CAAATTAGCT TACTGAATAC TGACCAGTTA  
1951 CTATAAATTT ATGATCTGGC TTTTGCACCC TGTTACAGTC TGCAGCATTA

FIGURE 19b (cont.)

36/66

2001	GTAGGTGACT	TCAACAATTG	GAATCCAAAT	GCAGATACTA	TGACCAGAGT
2051	ATGTCTACAG	CTTGGCAATT	TTCCACCTTT	GCTTCATAAC	TACTGATACA
2101	TCTATTTGTA	TTTATTTAGC	TGTTTGCACA	TTCCTTAAAG	TTGAGCCTCA
2151	ACTACATCAT	ATCAAAATGG	TATAATTTGT	CAGTGTCTTA	AGCTTCAGCC
2201	CAACATTCT	ACTGAATTTA	GTCCATCTTT	TTGAGATTGA	AAATGAGTAT
2251	ATTAAGGATG	AATGAATACG	TGCAACACTC	CCATCTGCAT	TATGTGTGCT
2301	TTTCCATCTA	CAATGAGCAT	ATTTCCATGC	TATCAGTGAA	GGTTTGCTCC
2351	TATTGATGCA	GATATTTGAT	ATGGTCTTTT	CAGGATGATT	ATGGTGTTTG
2401	GGAGATTTTC	CTCCCTAACA	ACGCTGATGG	ATCCTCAGCT	ATTCCTCATG
2451	GCTCACGTGT	AAAGGTAAGC	TGGCCAATTA	TTTAGTCGAG	GATGTAGCAT
2501	TTTCGAACTC	TGCCTACTAA	GGGTCCCTTT	TCCTCTCTGT	TTTTTAGATA
2551	CGGATGGATA	CTCCATCCGG	TGTGAAGGAT	TCAATTTCTG	CTTGGATCAA
2601	GTTCTCTGTG	CAGGCTCCAG	GTGAAATACC	TTTCAATGGC	ATATATTATG
2651	ATCCACCTGA	AGAGGTAAGT	ATCGATCTAC	ATTACATTAT	TAAATGAAAT
2701	TTCCAGTGTT	ACAGTTTTTT	AATACCCACT	TCTTACTGAC	ATGTGAGTCA
2751	AGACAATACT	TTTGAATTTG	GAAGTGACAT	ATGCATTAAT	TCACCTTCTA
2801	AGGGCTAAGG	GGCAACCAAC	CTTGGTGATG	TGTGTATGCT	TGTGTGTGAC
2851	ATAAGATCTT	ATAGCTCTTT	TATGTGTTCT	CTGTTGGTTA	GGATATTCCA
2901	TTTTGGCCTT	TTGTGACCAT	TTACTAAGGA	TATTTACATG	CAAATGCAGG
2951	AGAAGTATGT	CTTCCAACAT	CTCAACTAAA	CGACCAGAGT	CACTAAGGAT

FIGURE 19b (cont.)

3001 TTATGAATCA CACATTGGAA TGAGCAGCCC GGTATGTCAA TAAGTTATTT  
 3051 CACCTGTTTC TGGTCTGATG GTTTATTCTA TGGATTTTCT AGTTCTGTTA  
 3101 TGTACTGTTA ACATATTACA TGGTGCATTC ACTTGACAAC CTCGATTTTA  
 3151 TTTTCTAATG TCTTCATATT GGCAAGTGCA AAACTTTGCT TCCTCTTTGT  
 3201 CTGCTTGTTT TTTTGTCTTC TGTAAGATTT CCATTGCATT TGGAGGCAGT  
 3251 GGGCATGTGA AAGTCATATC TATTTTTTTT TTGTCAGAGC ATAGTTATAT  
 3301 ATTGTTGTTG CAATAGCTCG GTATAATGTA ACCATGTTAC TAGCTTAAGA  
 3351 TTTCCCACTT AGGATGTAAG AAATATTGCA TTGGAGCGTC TCCAGCAAGC  
 3401 CATTTCTTAC CTTATTAATG AGAGAGAGAC AAGGGGGGGG GGGGGGGGGG  
 3451 GGTTCCTTC ATTATTCTGC GAGCGATTCA AAAACTTCCA TTGTTCTGAG  
 3501 GTGTACGTAC TGCAGGGATC TCCCATTATG AAGAGGATAT AGTTAATTCT  
 3551 TTGTAACCTA CTTGGAAACT TGAGTCTTGA GGCATCGCTA ATATATACTA  
 3601 TCATCACAAT ACTTAGAGGA TGCATCTGAA nATTTTAGTG TGATCTTGCA  
 3651 CAGGAACCGA AGATAAATTC ATATGCTAAT TTTAGGGATG AGGTGTTGCC  
 3701 AAGAATTAAA AGGCTTGGAT ACAATGCAGT GCAGATAATG GCAATCCAGG  
 3751 AGCATTCATA CTATGCAAGC TTTGGGTATT CACACAATCC ATTTTTTTCT  
 3801 GTATACACnT CTTCACCCAT TTGGAGCTAT TACATCCTAA TGCTTCATGC  
 3851 ACATAAAATA TTTGGATATA ATCCTTTATT AGATATATAG TACAACTACA  
 3901 CTTAGTATTC TGAnnAAAnAA GATCATTTTA TTGTTGTTGG CTTGTTCCAG  
 3951 GTACCATGTT ACTAATTTTT TTGCACCAAG TAGCCGTTTT GGAACCTCAG

FIGURE 19b (cont.)

4001 AGGACTTAAA ATCCTTGATC GATAGAGCAC ATGAGCTTGG TTTGCTTGTT  
 4051 CTTATGGATA TTGTTTCATAG GTAATTAGTC CAATTTAATT TTAGCTGTTT  
 4101 TACTGTTTAT CTGGTATTCT AAAGGGAAAT TCAGGCAATT ATGATACATT  
 4151 GTCAAAAGCT AAGAGTGGCG AAAGTGAAAT GTCAAAATCT AGAGTGGCAT  
 4201 AAGGAAAATT GGCAAAAAC T AGAGTGGCAA AAATAAAATT TTCCCATCCT  
 4251 AAATGGCAGG GCCCTATCGC CGAATATTTT TCCATTCTAT ATAATTGTGC  
 4301 TACGTGACTT CTTTTTCTC AGATGTATTA AACCAGTTGG ACATGAAATG  
 4351 TATTTGGTAC ATGTAGTAAA CTGACAGTTC CATAGAATAT CGTTTTGTAA  
 4401 TGGCAACACA ATTTGATGCC ATAGATGTGG ATTGAGAAGT TCAGATGCTA  
 4451 TCAATAGAAT TAATCAACTG GCCATGTACT CGTGGCACTA CATATAGTTT  
 4501 GCAAGTTGGA AAAGTACAG CAATACCTCA CTGATAAGTG GCCAGGCCCC  
 4551 ATTTGAACAT ATTACTTAAA GTTCTTCATT TGTCCTAAGT CAACTTCTT  
 4601 TAAGTTTGAC CAAGTCTATT GGAAAATATA TCAACATCTA CAACACCAAA  
 4651 TTACTTTGAT CAGATTAACA ATTTTATTTT TATTATATTA GCACATCTTT  
 4701 GATGTTGTAG ATATCAGCAC ATTTTCTAT AGACTTGGTC AAATATAGAG  
 4751 AAGTTTGACT TAGGACAAAT CTAGAACTTC AATCAATTG GATCAGAGGG  
 4801 AACATCAAAT AATATAGATA GATGTCAACA CTTCAACAAA AAAATCAGAC  
 4851 CTTGTCACCA TATATGCATC AGACCATCTG TTTGCTTTAG CCACTTGCTT  
 4901 TCATATTTAT GTGTTTGATC CTAATCTACT TTCCTTCTA CTTGGTTTGG  
 4951 TTGATTCTAT TTCAGTTGCA TTGCTTCATC AATGATTTTG TGTACCCTGC

FIGURE 19b (cont.)



5001 AGTCATTTCGT CAAATAATAC CCTTGACGGT TTGAATGGTT TCGATGGCAC  
 5051 TGATACACAT TACTTCCACG GTGGTCCACG CGGCCATCAT TGGATGTGGG  
 5101 ATTCTCGTCT ATTCAACTAT GGGAGTTGGG AAGTATGTAG CTCTGACTTC  
 5151 TGTCACCATA TTTGGCTAAC TGTTCTGTGTT AATCTGTTCT TACACATGTT  
 5201 GATATTCTAT TCTTATGCAG GTATTGAGAT TCTTACTGTC AAACGCGAGA  
 5251 TGGTGGCTTG AAGAATATAA GTTTGATGGA TTTCGATTTG ATGGGGTGAC  
 5301 CTCCATGATG TATACTCACC ATGGATTACA AGTAAGTCAT CAAGTGGTTT  
 5351 CAGTAACTTT TTTAGGGCAC TGAAACAATT GCTATGCATC ATAACATGTA  
 5401 TCATGATCAG GACTTGTGCT ACGGAGTCTT AGATAGTTCC CTAGTATGCT  
 5451 TGTACAATTT TACCTGATGA GATCATGGAA GATTGGAAGT GATTATTATT  
 5501 TATTTTCTTT CTAAGTTTGT TTCTTGTTCT AGATGACATT TACTGGGAAC  
 5551 TATGGCGAAT ATTTTGGATT TGCTACTGAT GTTGATGCGG TAGTTTACTT  
 5601 GATGCTGGTC AACGATCTAA TTCATGGACT TTATCCTGAT GCTGTATCCA  
 5651 TTGGTGAAGA TGTAAGTGCT TACAGTATTT ATGATTTTTTA ACTAGTTAAG  
 5701 TAGTTTTTATT TTGGGGATCA GTCTGTTACA CTTTTTGTTA GGGGTAAAAT  
 5751 CTCTCTTTTC ATAACAATGC TAATTTATAC CTTGTATGAT AATGCATCAC  
 5801 TTAnGTAATT TGAAAAGTGC AAGGGCATT C AAGCTTACGA GCATATTTTT  
 5851 TGATGGCTGT AATTTATTTG ATAGTATGCT TGTTTGGGTT TTTCAATAAG  
 5901 TGGGAGTGTG TGAATAATGT TGTATTATTT ATTTAATTGC GGAAGAAATG  
 5951 GGCAACCTTG TCAATTGCTT CAGAAGGCTA ACTTTGATTC CATAAACGCT

FIGURE 19b (cont.)

40/66

6001 TTGGAAATGA GAGGCTATTC CCAAGGACAT GAATTATACT TCAGTGTGTT  
6051 CTGTACATGT ATTTGTAATA GTGGTTTAAC TTAAATTCCT GCACTGCTAT  
6101 GGAATCTCAC TGTATGTTGT nAGTGTACAC ATCCACAAAC AAGTAATCCT  
6151 GAGCTTTCAA CTCATGAGAA AATAnGAnGT CCGCTTCTGC CAGCATTAAC  
6201 TGTTACAGT TCTAATTTGT GTAAGTGTGA AATTGTTTCAAG GTCAGTGGAA  
6251 TGCCTACATT TTGCATCCCT GTTCCAGATG GTGGTGTGGG TTTTGACTAC  
6301 CGCCTGCATA TGGCTGTAGC AGATAAATGG ATTGAACTCC TCAAGTAAGT  
6351 GCAGGAATAT TGGTGATTAC ATGCGCACAA TGATCTAGAT TACATTTTCT  
6401 AAATGGTAAA AAGGAAAATA TGTATGTGAA TATCTAGACA TTTGCCTGTT  
6451 ATCAGCTTGA ATACGAGAAG TCAAATACAT GATTTAAATA GCAAATCTCG  
6501 GAAATGTAAT GGCTAGTGTC TTTATGCTGG GCAGTGTACA TTGCGCTGTA  
6551 GCAGGCCAGT CAACACAGTT AGCAATATTT TCAGAAACAA TATTATTTAT  
6601 ATCCGTATAT GAnGAAAGTT AGTATATAAA CTGTGGTCAT TAATTGTGTT  
6651 CACCTTTTGT CCTGTTTAAAG GATGGGCAGT AGGTAATAAA TTTAGCCAGA  
6701 TAAAATAAAT CGTTATTAGG TTTACAAAAG GAATATACAG GGTCATGTAG  
6751 CATATCTAGT TGTAATTAAT GAAAAGGCTG ACAAAGGCT CGGTAAAAAA  
6801 AACTTTATGA TGATCCAGAT AGATATGCAG GAACGCGACT AAAGCTCAAA  
6851 TACTTATTGC TACTACACAG CTGCCAATCT GTCATGATCT GTGTTCTGCT  
6901 TTGTGCTATT TAGATTTAAA TACTAACTCG ATACATTGGC AATAATAAAC  
6951 TTAAGTATTC AACCAATTTG GTGGATACCA GAnATTTCTG CCCTCTTGTT

FIGURE 19b (cont.)

7001 AGTAATGATG TGCTCCCTGC TGCTGTTCTC TGCCGTTACA AAAGCTGTTT  
7051 TCAGTTTTTT GCATCATTAT TTTTGTGTGT GAGTAGTTTA AGCATGTTTT  
7101 TTGAAGCTGT GAGCTGTTGG TACTTAATAC ATTCTTGGAA GTGTCCAAAT  
7151 ATGCTGCAGT GTAATTTAGC ATTTCTTTAA CACAGGCAA GTGACGAATC  
7201 TTGGAAAATG GCGGATATTG TGCACACCCT AACAAATAGA AGGTGGCTTG  
7251 AGAAGTGTGT AACTTATGCA GAAAGTCATG ATCAAGCACT AGTTGGTGAC  
7301 AAGACTATTG CATTCTGGTT GATGGATAAG GTACTAGCTG TTACTTTTGG  
7351 ACAAAGAAT TACTCCCTCC CGTTCCTAAA TATAAGTCTT TGTAGAGATT  
7401 CCACTATGGA CCACATAGTA TATAGATGCA TTTTAGAGTG TAGATTCACT  
7451 CATTTTGCTT CGTATGTAGT CCATAGTGAA ATCTCTACAG AGACTTATAT  
7501 TTAGGAACGG AGGGAGTACA TAATTGATTT GTCTCATCAG ATTGCTAGTG  
7551 TTTTCTTG TG ATAAAGATTG GCTGCCTCAC CCATCACCAG CTATTTCCCA  
7601 ACTGTTACTT GAGCAGAATT TGCTGAAAAC GTACCATGTG GTACTGTGGC  
7651 GGCTTG TGAA CTTTGACAGT TATGTTGCAA TTTTCTGTTC TTATTTATTT  
7701 GATTGCTTAT GTTACCGTTC ATTTGCTCAT TCCTTTCCGA GACCAGCCAA  
7751 AGTCACGTGT TAGCTGTGTG ATCTGTTATC TGAATCTTGA GCAAATTTTA  
7801 TTAATAGGCT AAAATCCAAC GAATTATTTG CTTGAATTTA AATATACAGA  
7851 CGTATAGTCA CCTGGCTCTT TCTTAGATGA TTACCATAGT GCCTGAAGGC  
7901 TGAAATAGTT TTGGTGTTTC TTGGATGCCG CCTAAAGGAG TGATTTTAT  
7951 TGGATAGATT CCTGGCCGAG TCTTCGTTAC AACATAACAT TTTGGAGATA

FIGURE 19b (cont.)

8001 TGCTTAGTAA CAGCTCTGGG AAGTTTGGTC ACAAGTCTGC ATCTACACGC  
8051 TCCTTGAGGT TTTATTATGG CGCCATCTTT GTAAC TAGTG GCACCTGTAA  
8101 GGAAACACAT TCAAAAGGAA ACGGTCACAT CATTCTAATC AGGACCACCA  
8151 TACTAAGAGC AAGATTCTGT TCCAATTTTA TGAGTTTTTG GGA CTCCAAA  
8201 GGGAACAAAA GTGTCTCATA TTGTGCTTAT AACTACAGTT GTTTTTATAC  
8251 CAGTGTAGTT TTATTCCAGG ACAGTTGATA CTTGGTACTG TGCTGTAAAT  
8301 TATTTATCCG ACATAGAACA GCATGAACAT ATCAAGCTCT CTTTGTGCAG  
8351 GATATGTATG ATTTTCATGGC TCTGGATAGG CTTCAACTCT TCGCATTGAT  
8401 CGTGGCATAG CATTACATAA AATGATCAGG CTTGTCACCA TGGGTTTAGG  
8451 TGGTGAAGGC TATCTTAACT TCATGGGAAA TGAGTTTGGG CATCCTGGTC  
8501 AGTCTTTACA ACATTATTGC ATTCTGCATG ATTGTGATTT ACTGTAATTT  
8551 GAACCATGCT TTTCTTTTAC ATTGTATGTA TTATGTAATC TGTTGCTTCC  
8601 AAGGAGGAAG TTAAC TTCTA TTTACTTGGC AGAATGGATA GATTTTCCAA  
8651 GAGGCCACAA AACTCTTCCA ACCGGCAAAG TTCTCCCCTG GAAATAACAA  
8701 TAGTTATGAT AAATGCCGCC GTAGATTGA TCTTGTAAGT TTTAGCTGTG  
8751 CTATTACATT CCCTCACTAG ATCG

FIGURE 19b (cont.)

# COMPARISON OF N-TERMINAL SEQUENCES OF SOLUBLE STARCH SYNTHASE

**GRYVAELSRGPAARP**      Deduced from wheat cDNA

**GPYVAELSPGPAAPP**      Wheat N-terminal

FIGURE 20a

1 TCTCCCACTC TTCTCTCCCC GCGCACACCG AGTCGGCACC GGCTCATCAC  
51 CCATCACCTC GGCCTCGGCC ACCGGCAAAC CCCCCGATCC GCTTTTGACG  
101 GCAGCGCACT AAAACCCCGG GGAGCGCGCC CCGCGGCAGC AGCAGCACCG  
151 CAGTGGGAGA GAGAGGCTTC GCCCGGGCCC GCACCGAGCG GGGCGATCCA  
201 CCGTCCGTGC GTCCGCACCT CCTCCGCCTC CTCCCCTGTC CCGCGCGCCC  
251 ACACCCATGG CGGCGACGGG CGTCGGCGCC GGGTGCCTCG CCCCCAGCGT  
301 CCGCCTGCGC GCCGATCCGG CGACGGCGGC CCGGGCGTCC GCCTGCGTCG  
351 TCCGCGCGCG GCTCCGGCGC TTGGCGCGGG GCCGCTACGT TGCCGAGCTC  
401 AGCAGGGAGG GCCCCGCGGC GCGCCCCGCG CAGCAGCAGC AACTGGCCCC  
451 GCCGCTCGTG CCAGGCTTCC TCGCGCCGCC GCCGCCCGCG CCGCCCCAGT  
501 CGCCGGCCCC GACGCAGCGG CCCCTGCCGG ACGCCGGCGT GGGGGAACTC  
551 GCGCCCCGACC TCCTGCTCGA AGGGATTGCT GAGGATTCCA TCGACAGCAT  
601 AATTGTGGCT GCAAGTGAGC AGGATTCTGA GATCATGGAT GCGAATGAGC  
651 AACCTCAAGC TAAAGTTACA CGTAGCATCG TGTTTGTGAC TGGTGAAGCT  
701 GCTCCTTATG CAAAGTCAGG GGGGCTGGGA GATGTTTGTG GTTCGTTACC  
751 AATTGCTCTT GCTGCTCGTG GTCACCGTGT GATGGTTGTA ATGCCAAGAT  
801 ACTTGAATGG GTCCTCTGAT AAAAATATG CAAAGGCATT ATACACTGGG  
851 AAGCACATTA AGATTCCATG CTTTGGGGGA TCACATGAAG TGACCTTTTT  
901 TCATGAGTAT AGAGACAACG TCGATTGGGT GTTTGTGAT CATCCGTCAT  
951 ATCATAGACC AGGAAGTTTA TATGGAGATA ATTTTGGTGC TTTTGGTGAT

FIGURE 20b

1001 AATCAGTTCA GATACACACT CCTTTGCTAT GCTGCATGCG AGGCCCCACT  
1051 AATCCTTGAA TTGGGAGGAT ATATTTATGG ACAGAATTGC ATGTTTGTTG  
1101 TGAACGATTG GCATGCCAGC CTTGTGCCAG TCCTTCTTGC TGCAAAATAT  
1151 AGACCATACG GTGTTTACAG AGATTCCCGC AGCACCCCTG TTATACATAA  
1201 TTTAGCACAT CAGGGTCTGG AGCCTGCAAG TACATATCCT GATCTGGGAT  
1251 TGCCACcTGA ATGGTATGGA GCTTTAGAAT GGGTATTTCC AGAATGGGCA  
1301 AGGAGGCATG CCCTTGACAA GGGTGAGGCA GTTAACTTTT TGAAAGGAGC  
1351 AGTCGTGACA GCAGATCGAA TTGTGACCGT CAGTCAGGGT TATTCATGGG  
1401 AGGTCACAAC TGCTGAAGGT GGACAGGGCC TCAATGAGCT CTTAAGCTCC  
1451 CGAAAAAGTG TATTGAATGG AATTGTAAAT GGAATTGACA TTAATGATTG  
1501 GAACCCCAAC ACAGACAAGT GTCTCCCTCA TCATTATTCT GTCGATGACC  
1551 TCTCTGGAAG GGCCAAATGT AAAGCTGAAT TGCAGAAGGA GCTGGGTTTA  
1601 CCTGTAAGGG AGGATGTTCC TCTGATTGGC TTTATTGGAA GACTGGATTA  
1651 CCAGAAAGGC ATTGATCTCA TTAAAATGGC CATTCCAGAG CTCATGAGGG  
1701 AGGACGTGCA GTTTGTCATG CTTGGATCTG GGGATCCAAT TTTTGAAGGC  
1751 TGGATGAGAT CTACCGAGTC GAGTTACAAG GATAAATTCC GTGGATGGGT  
1801 TGGATTTAGT GTTCCAGTTT CCCACAGAAT AACTGCAGGT TGCGATATAT  
1851 TGTTAATGCC ATCCAGGTTT GAACCTTGTG GTCTTAATCA GCTATATGCT  
1901 ATGCAATATG GTACAGTTCC TGTAAGTTCAT GGAAGTGGGG GCCTCCGAGA  
1951 CACAGTCGAG ACCTTCAACC CTTTGGGTGC AAAAGGAGAG GAGGGTACAG

FIGURE 20b (cont.)

2001 GGTGGGCGTT CTCACCGCTA ACCGTGGACA AGATGTTGTG GGCATTGCGA  
2051 ACCGCGATGT CGACATTCAG GGAGCACAAG CCGTCCTGGG AGGGGCTCAT  
2101 GAAGCGAGGC ATGACGAAAG ACCATACGTG GGACCATGCC GCCGAGCAGT  
2151 ACGAGCAGAT CTTCGAATGG GCCTTCGTGG ACCAACCCTA CFTCATGTAG  
2201 ACGGGGACTG GGGAGGTCGA AGCGCGGGTC TCCTTGAGCT CTGAAGACAT  
2251 GTTCCTCATC CTTCCGCGGC CCGGAAGGAT ACCCCTGTAC ATTGCGTTGT  
2301 CCTGCTACAG TAGAGTCGCA ATGCGCCTGC TTGCTTGGTC CGCCGGTTTG  
2351 AGAGTAGATG ACGGCTGTGC TGCTGCGGCG GTGACAGCTT CGGGTGGATG  
2401 ACAGTTACAG TTTTGGGGAA TAAGGAAGGG ATGTGCTGCA GGATGGTTAA  
2451 CAGCAAAGCA CCACTCAGAT GGCAGCCTCT CTGTCCGTGT TACAGCTGAA  
2501 ATCAGAAACC AACTGGTGAC TCTTTAGCCT TAGCGATTGT GAAGTTTGTT  
2551 GCATTCTGTG TATGTTGTCT TGTCCTTAGC TGACAAATAT TAGACCTGTT  
2601 GGAGAATTTT ATTTATCTTT GCTGCTGTTG TTTTGT TTT GTTAAAAAA  
2651 AAAAAAAAAA AA

FIGURE 20b (cont.)



1 GAGCTCCGAG AAnAGATTCC TATCATCGTC TTGGTGAGGT GAGGTTATGG  
 51 TTTCTTGTCA TGTGGGCAGA TTTGGTGCCA GATGCTTCAT ATCTATTCAA  
 101 GGGTTCAGCG GCAACAAC TG GCTCCAGA GCGATGGTCC TTAAGGGCAC  
 151 GTGCACGAAG ACTTCACGGC TGTTATCGAC AAGGTCAAGC CGGCTCCGAT  
 201 AGGGGAGCAG CGACAGCGGC GCGTCAACCG CTCGTTCTGG CGGCAGTAGT  
 251 GGTCGTTTCGG TGCTCTCGGA ACCTCGATGT AATTTTTATG ATTTTAGAGA  
 301 TGCTTTGTAc TTCcGATCGa TGAACCTCTGA TAATAGATAT CTcTTCTcTc  
 351 GCAAAAAAaG aGAGTTTTCA AcTGAAAACA AAaGaGTTTC AcTAGTTCTT  
 401 CTTTTAGAAA CAGAGTTTCA cTAGCAcTTT TTTTTCGcGAG AAGTcGAGTT  
 451 TCACtAAGTA cTAAaCCcAC GCAaTTATTC TCAAAAAAAA AACCCAcGcA  
 501 ACTGTcTGgA TcCATCTTCG TTTTTTCCCC GAGAATCGTC TGgATcCATT  
 551 TTCGTGTGCG AgGCATCCTC TCATTTTGcA cGgcCcAGcT cTcTTcTcGC  
 601 CGGcGTAcGc TGctAcATgT cGgcAcTCcA cGCAAACAAA AaGAaGCCCA  
 651 ACCGAAAACG cAcGcGCcTT TcCAgGcTCA ccACGGaAAA AAaTACcAcG  
 701 cGCcGcTcAC GAgCAAACCG TgACAACAGC CAGCCAGATA TGGCAACGGA  
 751 GGcACGGGCC GcACACAGCC AcTGAAAACC GCAGcTGcTC TTCCGTCCGT  
 801 CCGTCCcTCC GCCCGTCCGC gCcAcTCCAc TCGCCTTGCC CCAcTCCcAc  
 851 TCTTCTCTCC CCGCGCACAC CGAGTCGGCA CCGGCTCATC ACCCATCACy  
 901 TCGGCcTCGG CCACCGGCAA ACCCCCCGAT CCGCTTTTGC AGGCAGCGCA  
 951 CTAAAACCCC GGGGAGCGCG CCCCgCgg.C AGCAGCAGCA CCGCAGTGGG  
 1001 AGAGAGAGGC TTCGCCCCGG CCCGCACCGA GCGGGGCGAT CCACCGTCCG  
 1051 TGCGTCCGCA CCTCCTCCGC CTCTCCcCT GTCCCGCGCG CCCACACCCA

FIGURE 20c

1101 TGGCGGCGAC GGGCGTCGGC GCCGGGTGCC TCGCCCCCAG CGTCCGCcTG  
 1151 CGCGCCGATC CGGCGACGGC GGCCCGGGCG TCCGctTGCG TCGTCCGCGC  
 1201 GCGGCTCCGG CGcTTGGCGC GGGgCCGyTA CGTCGCCGAG CTCaGcAgGG  
 1251 AGGGCCCCGC GGcGCGCCCC GCGCAGCAGC AGCAACTGGC CCCGCCGCTC  
 1301 GTGCCAGGCT TCCTCGCGCC GCCGCCGCC GCGCCCCGCC AGTCGCCGGC  
 1351 CCCGACGCAG CCGCCCCcTGC CGGACgCCGG CgTGGGgGAA CTCGCGCCCg  
 1401 ACcTCcTGCT CGAAGGTAAA AAACAaggct gaatcCtcAg atcaCtcCGc  
 1451 gTcttcgTTt taccAaAtac ggtactGcga aGtgGtgcTg TATaTGtgaa  
 1501 gTtTcTgtcg aTtTcttcct gacggaTgtt cagtcgattc agtTgTATAT  
 1551 aTGtgAtacg ttcgtTgttc atcgatcgtA cAgaTttacc agCACactAg  
 1601 atAgAaatcG AgaccgaCGc GggcAgatca AtAgaTTTtT ctagaskTTT  
 1651 wwTkGrwtCG TGAGATGATT GATTGGGGTG GCGTGTGCAT ACGATAGCGG  
 1701 TGCACCGCCG ATGTATCGGG GCATGTGCAC GTGGTTGGGT CTCAGCAGAC  
 1751 ATATCACTAG ACTGGTATCG TAATTTACTA GTACTACTGG AAAGAGGACT  
 1801 AAAAAGGCTA GGCCAAGTGC ACGCATGTTG GGAACGTTGT TAAATTGATG  
 1851 AGTTTGTcCT TTGCTTGGGC TGGTATTATT ACCAAAAAAT GGTGTTAGTC  
 1901 CCTGTACTTA TTAATGGGaA AATCtTAACA TGACACTgGG GTTTATGAGT  
 1951 CTCCAATTGT ATATTCTCAG CACTCAACTG ATTTTACTGA TACTGTAGTG  
 2001 GAAATGACAC GTGAGCAcCC CCCTTCAAGG AATGCAATGC TTCTTTCTGT  
 2051 TTTAtATTAC AGGAACTAGA AGGAGCtTCC ACCTTTGAGT ACAGAAGTAC  
 2101 TCCCTCCGTT CCAAAATAGA TGA CTCAACT TTGTACTAAT TTTGTACTAT  
 2151 AGTTAGTACA AAGTTGAGTC ATCTATTTTA GAACGGAGGG AGTAGTATCG

FIGURE 20c (cont.)

2201 AAATTGAAGA CCCTTGTATT ACTGTCTTGT TTTTCAATGA AAATGGGAGG  
2251 CCCATGCAGT AAGTCACATG GGCACCTGGG AGGCTGGGAT CATGTGTGCT  
2301 TTGCAGAGTA CTAGACCCAG CTCACCCTCT GTTAGATTAC TTGTTGGGCT  
2351 GCTACTTTGT GTTTGCTGTG CAGTATATCA GACATCCTGA ATTTGGCATC  
2401 TAGCTGAGAA CAGAATGCAG GTTGCACCAT TCTTATTATT GCTAAACTGT  
2451 TGTCACGCAA TTTATAAAGA ATGTGATCTT CTGAGTATTA ATTAATCATG  
2501 TTCTGCTAAT ATCTGTCCTC GCTCTGGTGT TGACAAATAT ACCATATGAA  
2551 TATTTTCCAT TTTGCAACCA GGGATTGCTG AGGATTCCAT CGACAGCATA  
2601 ATCGTGGCTG CAAGTGAGCA GGATTCTGAG ATCATGGATG CGAATGAGCA  
2651 ACCTCAAGCT AAAGTTACAC GTAGCATCGT GTTTGTGACT GGTGAAGCTG  
2701 CTCCTTATGC AAAGTCAGGG GGGCTGGGAG ATGTTTGTGG TTCGTTACCA  
2751 ATTGCTCTTG CTGCTCGTGG TCACCGTGTG ATGGTTGTAA TGCCAAGATA  
2801 CTTGAATGGG TCCTCTGATA AAAACTATGC AAAGGCATTA TACACTGCGA  
2851 AGCACATTAA GATTCCATGC TTTGGGGGAT CACATGAAGT GACCTTTTTT  
2901 CATGAGTATA GAGACAACGT CGATTGGGTG GGTACACAAT CACCTTCTTA  
2951 TTCTCTGTTG AATTGTAGCA ACTGTTTATC CTTGTTTACA CTTCTTTTAG  
3001 CCCTGCAAAG ACATATGTGA TTTCCATACT TTTTGTATTAT TTCCCTTGTA  
3051 CTCTTGCTCA TGAAGGTCAA AATATCATAT ATCCATGGAA GTCATGCATG  
3101 TGCCTAGTAT TTTTGGTGTC GGTGCCTTTA ACTTTCAGGG ATTAATACGT  
3151 GGAATTTGAT AACTAAAGTT TATTTTATTG AAAAAAATTG TAGGTTGGCT  
3201 GAGCCCACAG CCACGCAGTG GCACCACTGC TTGCACATGA TTTTGCATTT  
3251 CTGTTTGCAC CGAGCACTTC ATGTGAATAA GGTGTAAAAT CATAAAGTAC

FIGURE 20c (cont.)

3301 CAATTTTATT CTGCCAATTG CACTTAAGAG TATATACATT TATCTTGGCC  
 3351 TCAATCATGG GAGTACTGTG CATTCAGTGC ACCATCATTG TTCTAAGGAG  
 3401 AAAATGTGGG TGCAAGGAAG ACACCTTTTGT CCCTTAATAA AAGGCAGGCA  
 3451 CTCTGTTGTC ATATAGATAG AAAGCAACAA ACTTATTTCA AAGAGCTAAC  
 3501 AATGGCAAAA GAACCAAAAA AAGCATGCTA AGGCGGTGAC ACCAAAAGGT  
 3551 GAGGGGGGCC TTGTGACTGA CAGCACCCCA AACTATTGCC ATTGTTTTAC  
 3601 TAAATGAAGA TCATTTTAGA AGCTCTCAGG AACTTCGAAA ACAGTGGCTT  
 3651 TCCGTCCACA GATCGTCTGT TAATATTTTT GTCCAGTGAT ACTTTTTTTG  
 3701 CTCCTTACAA GAGTGCCTAT GTTGACATAT ACATTGTTAA GTTGTTTATA  
 3751 AGTTTACTTC TTATTCTAAA CAGCAAGTGC CTAATGCTTG CATTTATTTT  
 3801 GGCTATTTAT TTTTATTCTC ATTTCAATCA ACACTTTTGT TCAGGTGTTT  
 3851 GTCGATCATC CGTCATATCA TAGACCAGGA AGTTTATATG GAGATAATTT  
 3901 TGGTGCTTTT GGTGATAATC AGGTACACTA CACTATACTA AGCTCCTAGT  
 3951 TGACTAAGTC GTAAGTTGTA CCTCCTCGCT GACCGGCTGC TCTATGTCGT  
 4001 GCAGTTCAGA TACACACTCC TTTGCTATGC TGCATGCGAg GCCCCTACTAA  
 4051 TCCTTGAATT GGGAGGATAT ATTTATGGAC AGAATTGCAT GTTTGTTGTG  
 4101 AACGATTGGC ATGCCAGCCT TGTGCCAGTG TACGTTGTTT GTGGATCTGA  
 4151 AAGTCCAATC CTTTATTCAT TCTCTGCTTT GCAGTGTGCC CATGTCTACA  
 4201 TTTCTTTTAT GCTTTTTTCA TGTCTGTTCT TATATTGCAT ATATGCTTAT  
 4251 GGAGTCTAAA AGTTACCGGA GGAATAACT CtTAAGGAtT TCCTCAATCA  
 4301 ATTATCtTTA GcTTAGTTA ACAtTTACTG TGGCAAACAT AATGTGtTTT  
 4351 GAGAtTTACA ArkTCAGAGA TTgCACTTCA CTAGtTCGTA gCTAAAtCyGA

FIGURE 20c (cont.)

4401 tGtTTTCCCC GAGaAAATGC CtAAAGCTTT gtGTCTTGAT gCAtTGATAG  
 4451 aAAAAGAgTt TATGTaCACT CCcaAAGAgG GGACCcaAAA TTaCaACAcC  
 4501 AcACCCctGA GaACtAgGcG CtGCCgGAAg AAgCGATgCa AGccCCAcTG  
 4551 CCCCTGCCTT AGCTCAAAGC CGGGCgTCAG cCTTGATTrT GTCAAGTAAG  
 4601 CTAGCAGTGC TAGATTGCGC AAGGTCGATT CGTCGAAGAT GACAGTGTTG  
 4651 CGCTGCTTCC AAATCCACCA AACTATGAGC ATGATCACTG GAGAAGTACC  
 4701 TTTTCTCGCG GCTGAGGGGG TGGACTGGTG GTCTGCTGCT GCCAGTTTTC  
 4751 AGATAATCTG AAAAATGCAT GTTTTGATGA TTTTAGTATC TTGCGGACCC  
 4801 TGGGTACCAC CTAAGCTTTC ACACAGTAAT TTGCAGTTAC ACcTATAAAA  
 4851 GTAACGGTCA TGATATGCAT GTGTTTTGGG TAGATCATGG TGCATGCATT  
 4901 TTAGGAATTA GGACATGCCA GAACCACGTG AGGCTTATGG GGCAATTcAT  
 4951 TTGTTCCATT ATACGAGTCA TGAATATGGT TCAGCATGTT TGGACGCTAC  
 5001 TTGTTTGGGG CAATTTcAGA TGGTGAATTG TAGCTGCTTG ATGTTGGCTA  
 5051 GCTGGCTTAT TTTGTACAAG TATCGATGTT AGATGCATAT TTCCTTTTGT  
 5101 TCTTGTGCTG TTTGCCATGT TGTATTCCCC TTTTCTGTcG CCAGTGTTGC  
 5151 ATGTTAAATT GGTTTTcATT ACATAATCAA CTTTGTTGCT GACATCAGTC  
 5201 ATTTTTATTC AGCCTTCTTG CTGCAAAATA TAGACCATAC GGTGTTTACA  
 5251 GAGATTCCCG CAGCACCCCTT GTTATACATA ATTTAGCACA TCAGGTTTGG  
 5301 GTCTATCACC TTTCATTATC CGTACATGGC TTTGTAAGTC GGTTcACACG  
 5351 TATCGTCATA CTGTATGTTA TTTCAATGTC ATTAgGGTGT GGAGCCTGCA  
 5401 AGTACATATC CTGATCTGGG ATTGCCACCT GAATGGTATG GAGCTtTAGA  
 5451 ATGGGTATTT CCAGAATGGG CAAGGAGGCA TGCCCTTGAC AAGGGTGAGG

FIGURE 20c (cont.)

5501 CAGTTAACTT TTTGAAAGGA GCAGTTGTGA CAGCAGATCG AATTGTGACC  
5551 GTCAGTCAGG TGAAATACTC AATACTTCTC TTTTTTCTTT GCGGGATGTT  
5601 CTTCAGTTCA ATTGCCCTGT CTTTCACCCA ATTAAGAAAT GATTTAATCT  
5651 TTTGTTTCTA GGGTTATTCA TGGGAGGTCA CAACTGCTGA AGGTGGACAG  
5701 GGCcTCAATG AGCTCTTAAG CTCCCGAAAA AGTGTA tTGA ATGGTAACTA  
5751 TA tTTGAATC CACTTATCTT C . TTCTGAAA CATATTTACA GAAATAGATG  
5801 GATGGGTTGC AAGAATAAAT TCAGTTTGCT CTTTCGGTAT GAAGGAATTG  
5851 TAAATGGAAT TGACATTAAT GATTGGAACC CCACCACAGA CAAGTGTCTC  
5901 CCTCATCATT ATTCTGTCTG TGACCTCTCT GGAAAGGTGT GTGGATAGTA  
5951 CCcTATATAA TAACATGTAT ATCTGATC . T AGTACTTTCT TTTTCTTTGC  
6001 TAGTTTGCTT CCCATGATGT TCTCACTAAC TAATCCTATG TGGTTTGGCA  
6051 TACTTGTCAG GCCAAATGTA AAGCTGAATT GCAGAAGGAG CTGGGTTTAC  
6101 CTGTAAGGGA gGATGTTCcT CTGGTTaGAT ACAAACCCcT aAGATATaTA  
6151 TtTtTTAAAT CCCTAAAAAA AAcTTGCCGA TCATCTCaTT AGCTTGATTC  
6201 ACAGATTGGC TtTATTGGAA GACTGGATTA CCAGAAAGGC ATTGATCTCA  
6251 TTAAAATGGC CATTCCAGAG CTC

FIGURE 20c (cont.)

1 MAATGVGAGC LAPSVRLRAD PATAARASAC VVRARLRRLLA RGRYVAELSR  
51 EGPAARPAQQ QQLAPPLVPG FLAPPPPPA QSPAPTQPPL PDAGVGELAP  
101 DLLLEGIAED SIDSIIAAS EQDSEIMDAN EQPQAKVTRS IVFVTGEAAP  
151 YAKSGGLGDV CGSLPIALAA RGHRVMVMP RYLNGSSDKN YAKALYTGKH  
201 IKIPCFGGSH EVTFFHEYRD NVDWVFVDHP SYHRPGSLYG DNFGAFGDNQ  
251 FRYTLLCYAA CEAPLILELG GYIYGQNCMF VVNDWHASLV PVLLAAKYRP  
301 YGVYRDSRST LVIHNLAHQG LEPASTYPDL GLPPEWYGAL EWFPEWARR  
351 HALDKGEAVN FLKGAVVTAD RIVTVSQGYS WEVTTAEGGQ GLNELLSSRK  
401 SVLNGIVNGI DINDWNPTD KCLPHHYSVD DLSGKAKCKA ELQKELGLPV  
451 REDVPLIGFI GRLDYQKGID LIKMAIPELM REDVQFVMLG SGDPIFEGWM  
501 RSTESSYKDK FRGWVGFSVP VSHRITAGCD ILLMPSRFEP CGLNQLYAMQ  
551 YGTVPVVHGT GGLRDTVETF NPFQAKGEEG TGWAFSPLTV DKMLWALRTA  
601 MSTFREHKPS WEGLMKRGMT KDHTWDHAAE QYEQIFEWAF VDQPYVM\*

FIGURE 21

# Soluble Starch Synthase Genomic Clones

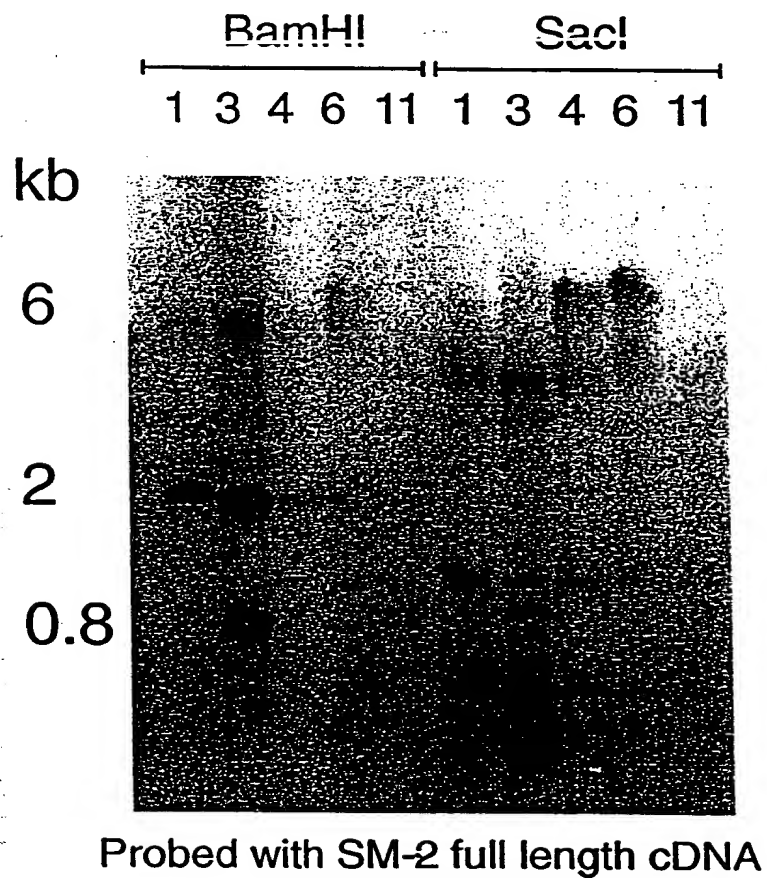


FIGURE 22



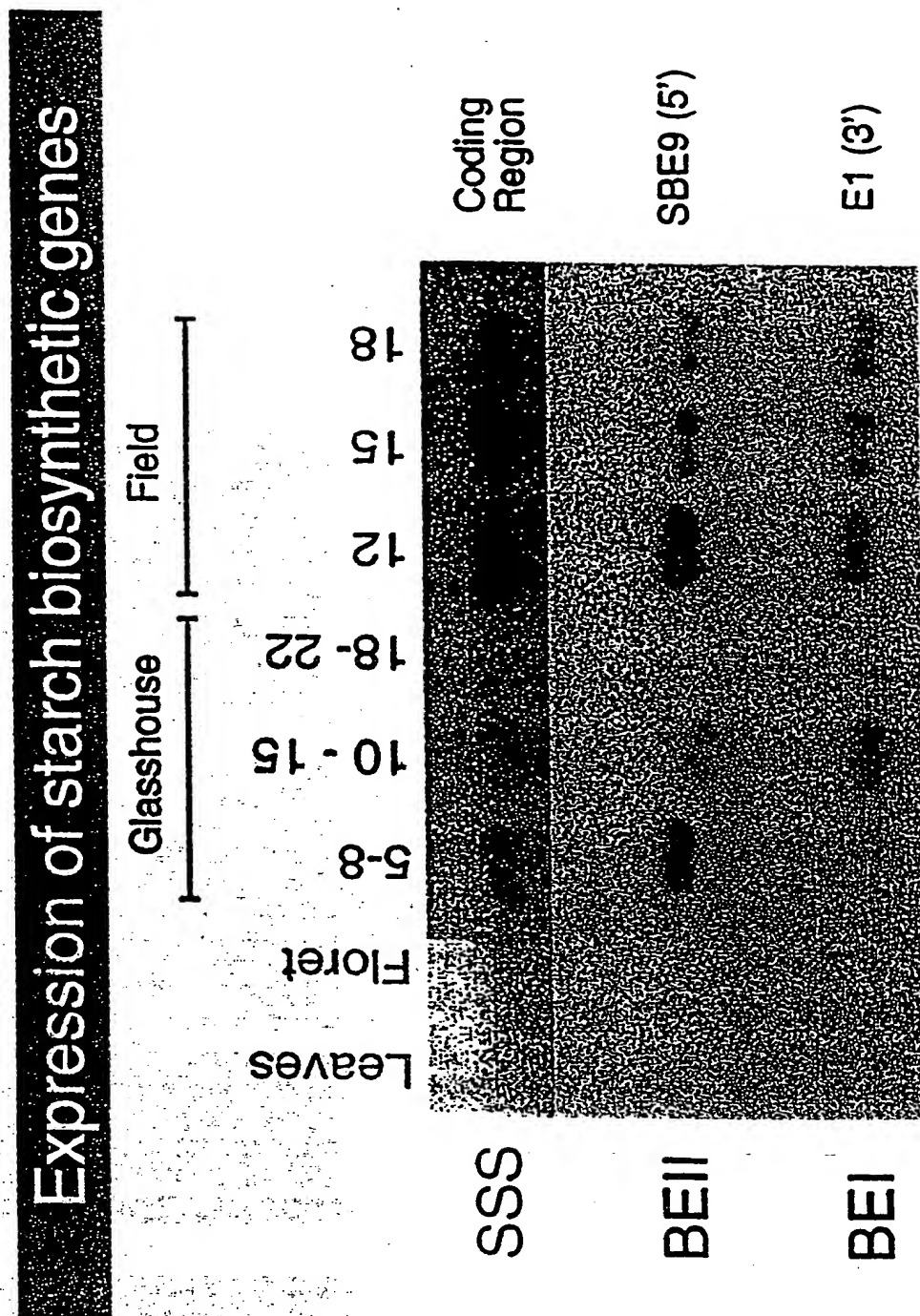
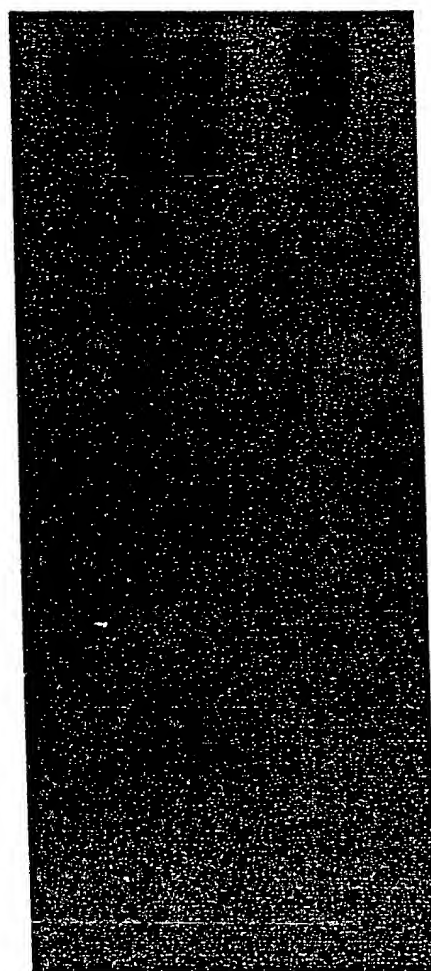


FIGURE 23

56/66

Wheat DNA probed with  
Soluble Starch Synthase

N7AT7D  
N7DT7B  
N7BT7A



8 kb

1.5 kb

FIGURE 24

1 GAGCTCCGAG AAnAGATTCC TATCATCGTC TTGGTGAGGT GAGGTTATGG  
 51 TTTCTTGTCa TGTGGGCAGA TTTGGTGCCA GATGCTTCAT ATCTATTCAA  
 101 GGGTTCAGCG GCAACAACtG CGGCTCCAGA GCGATGGTCC TTAAGGGCAC  
 151 GTGCACGAAG ACTTCACGGC TGTTATCGAC AAGGTCAAGC CGGCTCCGAT  
 201 AGGGGAGCAG CGACAGCGGC GCGTCAACCG CTCGTTCTGG CGGCAGTAGT  
 251 GGTCGTTTCGg TGCTCTCGGA ACCTCGATGT AATTTTTATG ATTTTAGAGA  
 301 TGCTTTGTAc TTCcGATCGa TGAACtCTGA TAATAGATAT CTcTTCTcTc  
 351 GCAAAAAAaG aGAGTTTTCA AcTGAAACA AAaGaGTTTC AcTAGTTCTT  
 401 CTTTTAGAAA CAGAGTTTTCA cTAGCAcTTT TTTTtGcGAG AAGTcGAGTT  
 451 TCActAAGTA cTAAaCCCAC GCAaTTATTC TCAAAAAAAA AACCcAcGcA  
 501 ACTGTcTGgA TcCATCTTCG TTTTTTCCCC GAGAATCGTC TGgATcCATT  
 551 TTCGTGTGCG AgGCATCCTC TCATTTTGcA cGgcCcAGcT cTcTTcTcGC  
 601 CGGcGTAcGc TGctAcATgT cGgcAcTCcA cGCAACAAA AaGAaGCCCA  
 651 ACCGAAAACg cAcGcGCcTT TcCAGGcTCA ccACGGaAAA AAaTACcAcG  
 701 cGcCgCtAcAC GAgCAAACCG TgACAACAGC CAGCCAGATA TGGCAACGGA  
 751 GGcACGGGCC GcACACAGCC AcTGAAAACC GCAGcTGcTC TTCCGTCCGT  
 801 CCGTCCcTCC GCCCGTCCGC gCcAcTCCAc TCGCCTTGCC CCAcTCCcAc  
 851 TCTTCTCTCC CCGCGCACAC CGAGTCGGCA CCGGCTCATC ACCCATCACy  
 901 TCGGCcTCGG CCACCGGCAA ACCCCCCGAT CCGCTTTTGC AGGCAGCGCA  
 951 CTAAAACCCC GGGGAGCGCG CCCCGcgg.C AGCAGCAGCA CCGCAGTGGG  
 1001 AGAGAGAGGC TTCGCCCCGG CCCGCACCGA GCGGGGCGAT CCACCGTCCG  
 1051 TGCGTCCGCA CCTCCTCCGC CTCCTCCCCT GTCCCGCGCG CCCACACCCA  
 1101 TGG

FIGURE 25

# Comparison of Wheat and Rice Soluble Starch Synthase Genomic DNA Sequences

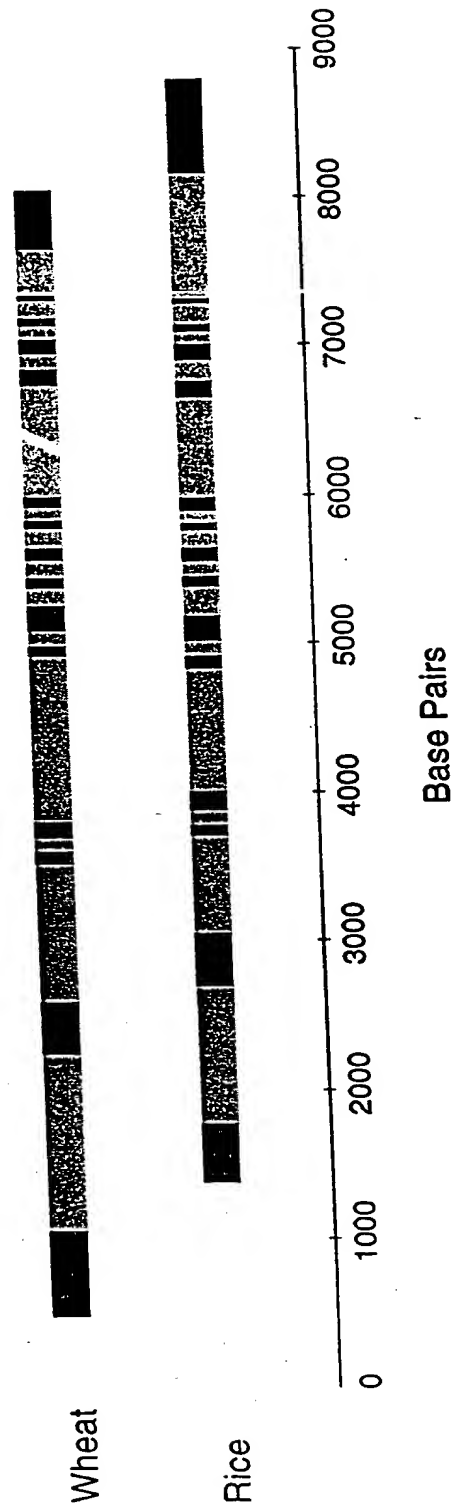


FIGURE 26

```

      80 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      ATACTACATACTATGCTTGCACCCAGGACACTTTTATAACTATTCTGGCTGTGGGA 139
      TATGATGATATATACGAACGTGGGTTCCCTGTGAAAATATTGATAAGACCGACACCCCT
      a      T T Y Y M L A P K G H F Y N Y S G C G N -
      b      I L H T I C L H P R D T F I T I L A V G -
      c      Y Y I L Y A C T Q G T L L * L F W L W E -

      140 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      ATACCTTCAACTGTAATCATCCTGTGGTTCGTCAATTCATTGTAGATTGTTAAGATACT 199
      TATGGAAGTTGACATTAGTAGGACACCAAGCAGTTAAGTAACATCTAACAAATTCATGA
      a      T F N C N H P V V R Q F I V D C L R Y W -
      b      I P S T V I I L W F V N S L * I V * D T -
      c      Y L Q L * S S C G S S I H C R L F K I L -

      200 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
      GGGTGACGGAAATGCATGTTGATGGTTTTCGTTTGGACCTT 240
      CCCACTGCCCTTTACGTACAACACTACCAAAAGCAAAACTGGAA
      a      V T E M H V D G F R F D L -
      b      G * R K C M L M V F V L T -
      c      G D G N A C * W F S F * P -

      Enzymes that do cut:
      NONE
      Enzymes that do not cut:
      EcoRI

```

FIGURE 27a

Comparison of Wheat Debranching Enzyme-I (WDBE-I) PCR fragment with maize *Sugary-1* DNA sequence

SUGARY.DNA	1098	1107	1117	1127	1137	1147	1157
		TGAGGTGATCATGGATGTTGCTTCAATCATACAGCTGAAGGTAATGAGAAGGCCCAAT					
WHEAT1.DNA							
		....GTGATCATGGATGTTGCTTCAACCATACAGCTGAGGTAATGAGAATGGTCCAAT					
	-3	6	16	26	36	46	56
FILE NAME	1158	1167	1177	1187	1197	1207	1217
SUGARY.DNA		ATTATCCTTTAGGGGATAGATAATAGTACTACTACATGCTTGACCTAAGGGAGGTT					
WHEAT1.DNA		ATTATCATTAGGGGGTTCGATAATACTACTACTATATGCTTGACCCCAAGGACACTT					
	57	66	76	86	96	106	116
FILE NAME	1218	1227	1237	1247	1257	1267	1277
SUGARY.DNA		TTATAATTATTCTGGTTGGGAAATACCTTCAATTGTAATCATCCTGTAGTCCGTGAATT					
WHEAT1.DNA		TTATAACTATTCTGGCTGGGNATACCTTCAACTGTAATCATCCTGTGTTTCGTCAATT					
	117	126	136	146	156	166	176
FILE NAME	1278	1287	1297	1307	1317	1327	1337
SUGARY.DNA		TATAGTGGATTGCTTGAGATACTGGGTAACAGAAATGCATGTTGATGGTTTCGTTTGA					
WHEAT1.DNA		CATTGTAGATTGTTTAAGNTACTGGGTGACGGAATGCATGTTGNTGGTTTCGTTTGA					
	177	186	196	206	216	226	236
FILE NAME	1338	1347	1357				
SUGARY.DNA		CCTTGCATCTACT-G...					
WHEAT1.DNA		CCTTGCATCTN--CTTAA					
	237	246	256				

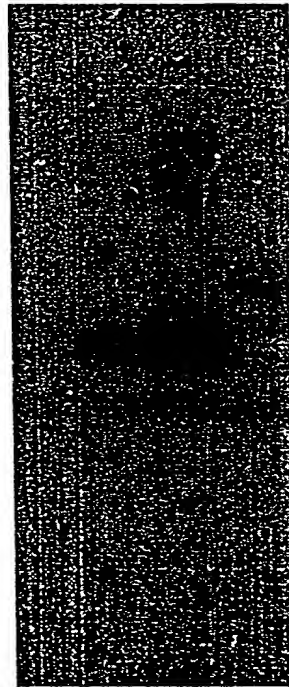
MATCHING PERCENTAGE  
TOTAL WINDOW 84% ( 219/ 260)  
ALIGNMENT WINDOW 86% ( 219/ 253)

FIGURE 27b

61/66

Southern blot of *T. tauschii*  
Genomic DNA

1X 3X



BamHI Digest

*T. tauschii* Genomic DNA Probed  
With The Wheat Debranching Enzyme  
PCR Product

FIGURE 28

Sequences of Primers which Direct PCR amplification of WSBEIL-D1 introns

Intron	Forward primer	Forward primer Seq	Reverse primer	Reverse primer Seq	Predicted Length of Product
1	sr854.1180F	CTG GCT GAC TCA ATC ACT ACG	WSBE9E2R	GGC ACG AGT GTG TGT ACC TGT A	601
2	WBE2E1F	CGT CGC TGC TCC TCA GGA AG	WBE2E2R	CAG GAC CTT CCC TGG AGA GG	401
3	WBE2E2F	CGC AAC CTG AAG AAT TAC AG	sr866F	TAT CTT CAG GTA TCT ACA GC	309
4	WBE2E3F	ATT TTC GGA GCC ATC TTG AC	WBE2E4R2	ATG CTT CCA ATC CAC CTT CA	>450
5	WBE2E4F	TCG TGG TTA TGA AAA GCT TGG	WBE2E5R	GAG CCC ATT CTC GGT AAG TGA	234
6	sr913F	ATC ACT TAC CGA GAA TGG G	WBE2E6R	CTG CAT TTG GAT TCC AAT TG	232
7	WBE2E6F	ACA ATT GGA ATC CAA ATG CA	WBE2E7R	GGG AGG AAA ATC TCC CAA AC	402
8	WBE2E7F	AGC TAT TCC TCA TGG CTC AC	sr915F	CCA TTG AAA GGT ATT TCA CC	203
9	WBE2E8F	TGC AGG CTC CAG GTG AAA TA	sr912F	TAA CTT ATT GAC ATA CC3 G	439



63/66

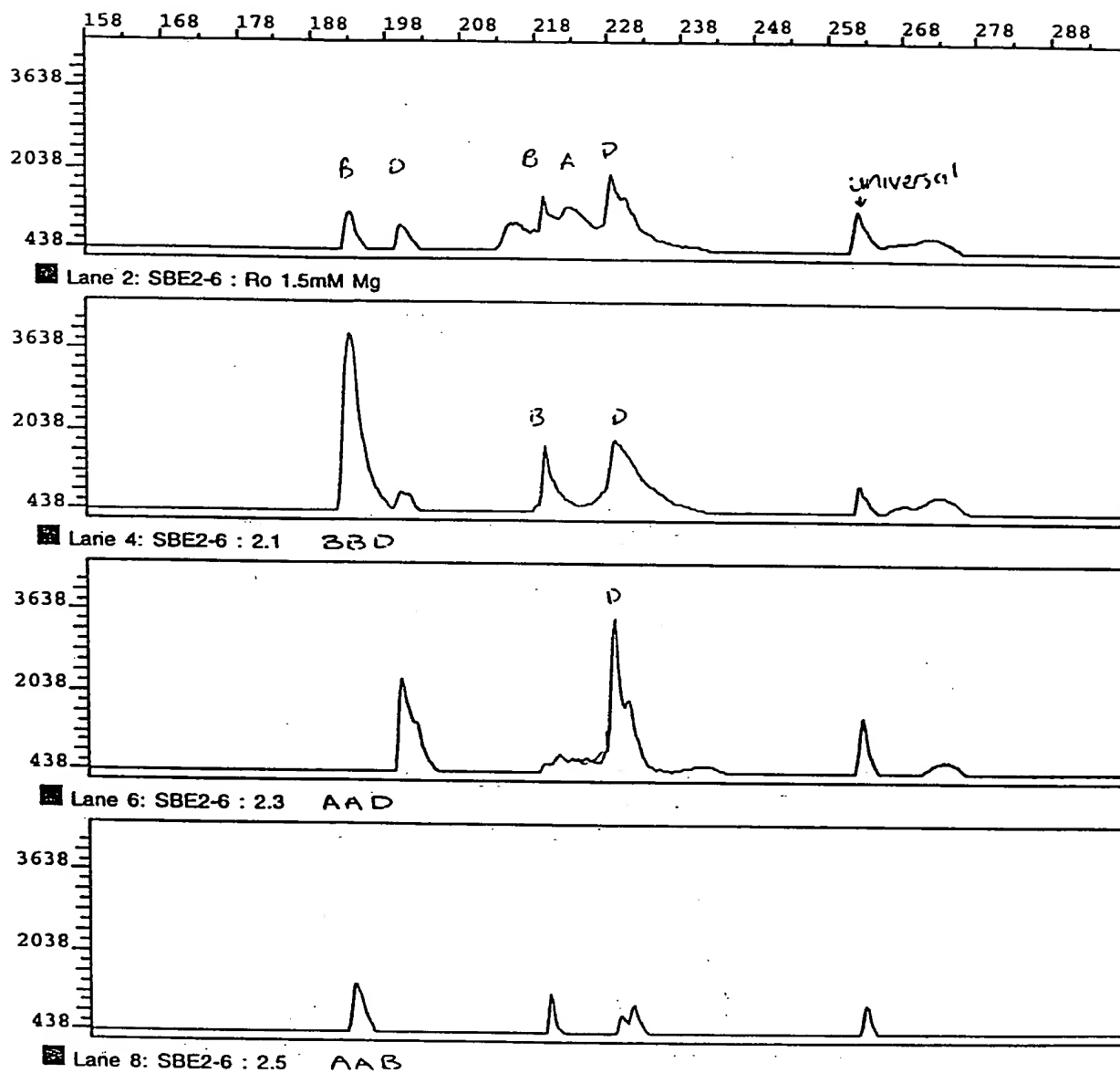


FIGURE 30

64/66

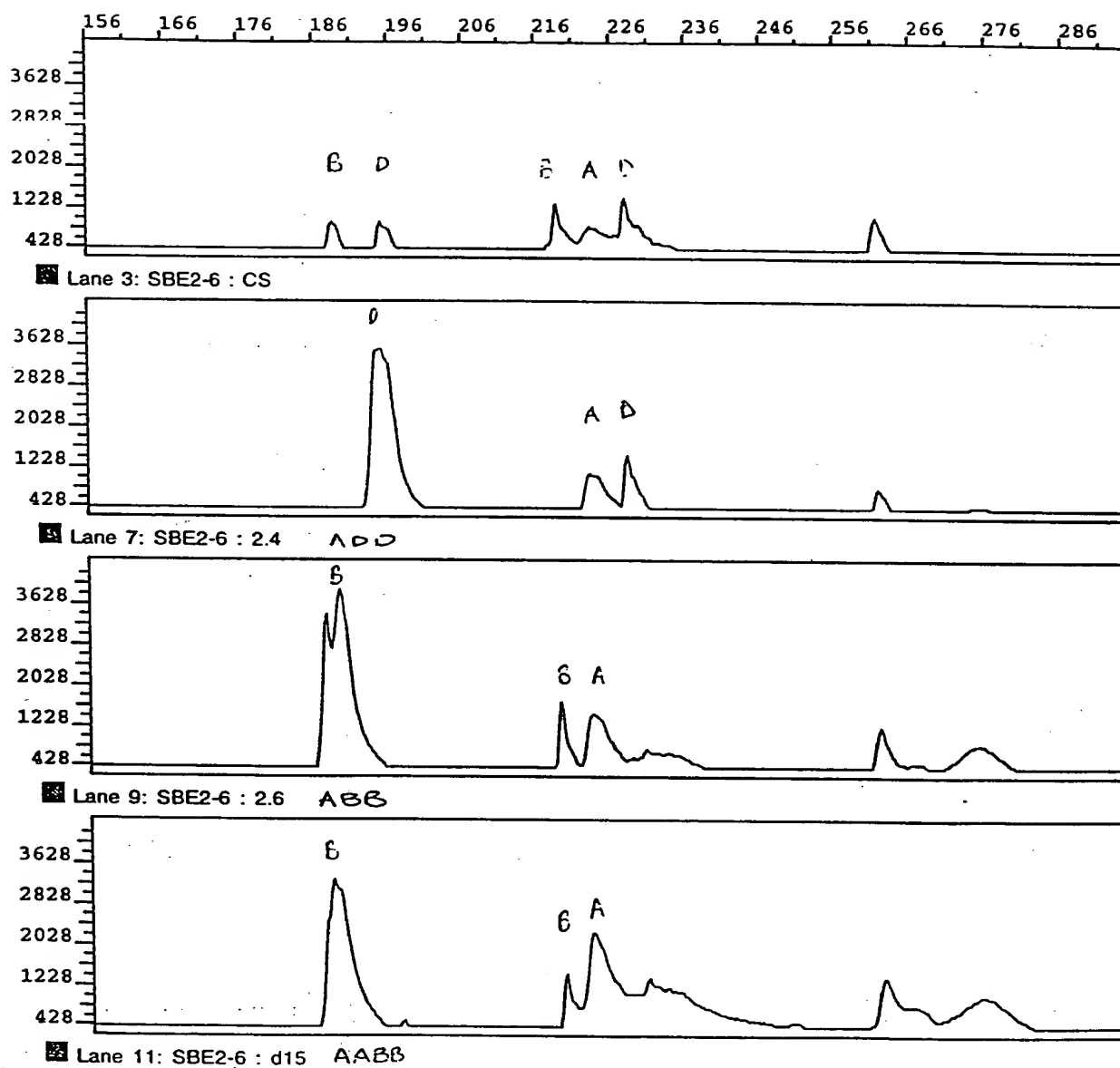


FIGURE 30 (cont.)

65/66

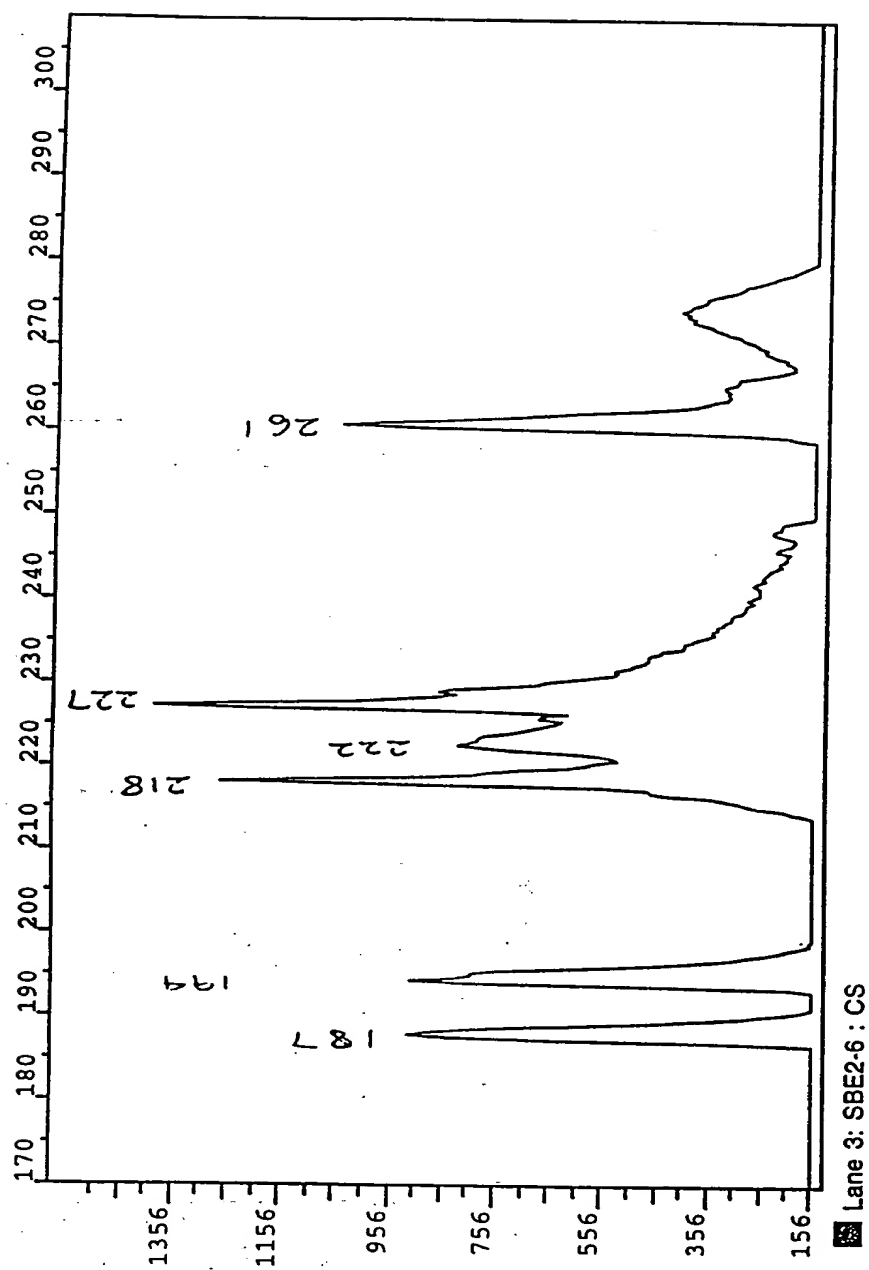


FIGURE 31a

66/66

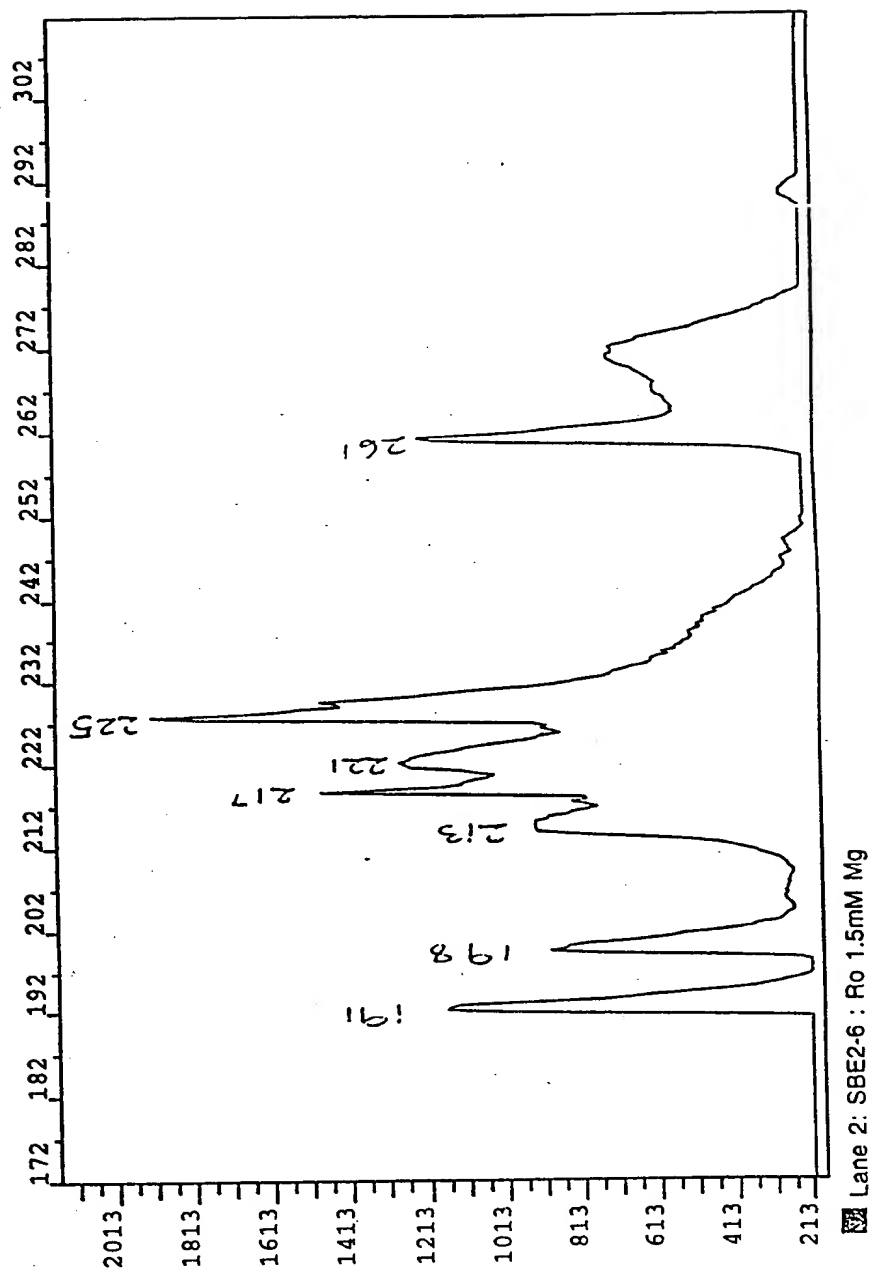


FIGURE 31b